

RISK FACTORS FOR OVARIAN CANCER
BY HISTOLOGICAL PHENOTYPE
IN THE MILLION WOMEN STUDY



Kezia Gaitskell
Green Templeton College
University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2016

Abstract

Risk factors for ovarian cancer by histological phenotype in the Million Women Study

Kezia Gaitskell, Green Templeton College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy, Trinity Term 2016

There is growing evidence that the different histological types of ovarian cancer have diverse origins. Many high-grade serous tumours (the most common type) are hypothesised to arise from the fallopian tubes, while endometrioid and clear cell tumours are hypothesised to develop from endometriosis, and the origins of mucinous tumours are uncertain. If these hypotheses are true, then the population-level risk factors for the ovarian cancer histotypes should likewise vary – but few epidemiological studies have sufficient cases to explore this. This thesis investigates the association between ovarian cancer and various exposures (including tubal ligation and reproductive factors), in a cohort of 1.3 million women, with 8,000 incident ovarian cancers. Participants completed a questionnaire at recruitment, and were followed up for routinely collected information on cancers and deaths using national registries. Cox proportional hazards models were used to estimate adjusted relative risks of ovarian cancer in women by different exposures. Tubal ligation was associated with reduced risks of cancers of the ovary, and also of the peritoneum and fallopian tube; parity and breastfeeding were also associated with reduced risks of ovarian cancer. These associations with ovarian cancer were highly heterogeneous between the different histological tumour types, consistent with hypotheses of their distinct origins. In particular, risks of endometrioid and clear cell tumours were substantially reduced amongst women with tubal ligation, but substantially increased amongst nulliparous women. For high-grade serous carcinomas, there was a smaller but significant reduction in risk with tubal ligation, and no significant association with parity. I also describe some methodological studies I set up for future tissue-based work: developing techniques in image analysis, and a pilot study of obtaining linked pathology reports and tissue samples. The population-level associations demonstrated by histotype are largely consistent with the new hypotheses of the different origins of ovarian cancer histotypes.

Acknowledgements

The work described in this thesis is only possible because of the 1.3 million women who participated in the Million Women Study (MWS), and the many researchers, data management and data entry staff, information technology and support staff of the Cancer Epidemiology Unit at the University of Oxford, past and present, who have worked on this study for more than two decades.

I am most grateful to my supervisors, Professor Dame Valerie Beral and Professor Ahmed Ahmed, for the opportunity to study and work in their research groups, and for their unfailing academic support, guidance and advice. I have also benefited immensely from discussions with other colleagues. I am particularly indebted to Kirstin Pirie and Isobel Barnes for help with STATA statistical programming (also acknowledged separately elsewhere). I would also like to thank Hayley Abbiss, Matt Arnold, Krys Baker, Angela Balkwill, Anna Brown, Benjamin Cairns, Dexter Canoy, Andrew Chadwick, Kate Coffey, Dave Ewart, Toral Gathani, Jane Green, Lynden Guiver, Wenji Guo, Carol Hermon, Sau Wan Kan, Seamus Kent, Tim Key, Ai Seon Kuan, Alison Price, Gillian Reeves, Julie Schmidt, Keith Shaw, Karl Smith Byrne, Ruth Travis, Lyndsey Trickett, and Ellie Watts.

I am also very grateful to Cancer Research UK (CRUK), and to the CRUK Oxford Centre in particular, not only for the generous financial support they have provided, but also for the opportunities for training and career support. I would also like to acknowledge the support of the Nuffield Department of Population Health, and of Green Templeton College.

Finally, I would like to thank my friends and family, who in the past three years have tolerated hearing considerably more about the histopathology and epidemiology of ovarian cancer than they perhaps desired or anticipated.

CONTENTS

List of Tables	ix
List of Figures	xi
List of Abbreviations	xv
The role of the author in this thesis	xvii
1 Thesis Introduction	1
1.1 Rationale	1
1.2 Research Aims	3
1.3 Outline of the thesis	3
1.4 Publications relating to this thesis	6
2 Ovarian cancer pathology	7
2.1 Introduction	7
2.2 New hypotheses regarding the origins of serous tumours	9
2.3 New hypotheses regarding the origins of Endometrioid and Clear Cell ovarian cancer	10
2.4 Hypotheses regarding the origins of mucinous tumours	12
2.5 Implications for this thesis	12

3	Ovarian cancer epidemiology	13
3.1	Ovarian cancer incidence and mortality	13
3.2	Randomised trials of screening	15
3.2.1	PLCO	15
3.2.2	UKCTOCS	16
3.3	Genetic risk	16
3.3.1	BRCA	16
3.3.2	Other genetic risks	17
3.4	Findings from the Collaborative Group on Epidemiological Studies of Ovarian Cancer	18
3.4.1	Heterogeneity by histological type	19
3.4.2	Retrospective vs prospective studies	21
3.5	Associations not yet reported on by the Collaboration	23
3.6	Mechanisms	26
4	The Million Women Study & Methods	28
4.1	Introduction	28
4.2	Design and Methods	28
4.3	Recruitment	29
4.3.1	Data collection	29
4.3.2	Data entry and storage	30
4.4	Ethics	30
4.5	Follow up	31
4.6	Exclusions	32
4.7	Outcome of incident ovarian cancer	32

4.7.1	Ovarian cancer by histological type	33
4.8	Statistical analyses	34
4.8.1	Descriptive statistics	34
4.8.2	Multivariate regression models	34
4.8.3	Plotting of figures	36
5	Tubal ligation and risk of ovarian cancer histotypes	41
5.1	Introduction	41
5.1.1	Tubal ligation: trends with age and time	41
5.1.2	Tubal ligation: associations with ovarian cancer	43
5.2	Aim	43
5.3	Methods	44
5.3.1	Study design, data collection and follow-up	44
5.3.1.1	Exposure variables	45
5.3.1.2	Outcome	45
5.3.2	Statistical analyses	46
5.4	Results	49
5.4.1	Variation by histotype	50
5.4.2	Timing of tubal ligation	51
5.4.3	Exploring possible effect modification	52
5.5	Discussion	57
5.5.1	Mechanisms	58
5.5.2	Strengths and weaknesses	59
5.5.3	Conclusion	61

6	Tubal ligation and risk of 26 site-specific cancers	63
6.1	Introduction	63
6.2	Aim	64
6.3	Methods	64
6.3.1	Study design, data collection and follow-up	64
6.3.1.1	Exposure variable	65
6.3.1.2	Outcomes	65
6.3.2	Statistical analysis	65
6.4	Results	68
6.5	Discussion	73
6.5.1	Tubal ligation and risk of ovarian, peritoneal and fallopian tube cancers	73
6.5.2	Tubal ligation and risk of other cancers	76
6.5.3	Strengths and weaknesses of this study	78
6.5.4	Conclusion	79
7	Reproductive factors and risk of ovarian cancer histotypes	80
7.1	Introduction	80
7.2	Aim	82
7.3	Methods	82
7.3.1	Study design, data collection and follow-up	82
7.3.2	Exposure variables	83
7.3.3	Outcome	83
7.3.4	Statistical analyses	84
7.4	Results	88

7.4.1	Parity and risk of ovarian cancer	88
7.4.2	Breastfeeding and risk of ovarian cancer	90
7.5	Discussion	98
7.5.1	Mechanisms	99
7.5.2	Strengths and weaknesses	100
7.5.3	Conclusion	100
8	Image analysis of the fallopian tubes: macro development and validation	102
8.1	Introduction	102
8.1.1	Project investigating SOX2 expression in the fallopian tubes of ovarian cancer patients	103
8.1.2	Existing methods of assessing immunohistochemical staining in the fallopian tube epithelium	104
8.2	Aim	105
8.3	Methods	105
8.3.1	Tissue samples and slides	105
8.3.2	Development of a macro to identify the fallopian tube epithelium .	107
8.3.3	Validation of the macro	108
8.3.3.1	Phase I	108
8.3.3.2	Phase II	109
8.4	Results	110
8.4.1	Phase I	110
8.4.1.1	Agreement	110
8.4.1.2	Timing	112

CONTENTS

8.4.2	Phase II	113
8.4.2.1	Agreement	113
8.4.2.2	Investigating outliers	115
8.4.3	Comparing inter-observer vs inter-method variation	122
8.5	Discussion	123
9	MWS Ovarian Cancer Tissue Pilot Study	124
9.1	Introduction	124
9.2	Aims	125
9.3	Methods	126
9.3.1	Initial planning and ethical approval	126
9.3.2	Research approval from the Oxford Centre for Histopathology Research (OCHRe)	126
9.3.3	Development of an initial protocol for tissue processing	127
9.3.4	Identification of participants and data transfer and linkage	127
9.3.5	Review of histopathology reports and identification of potential specimens	128
9.4	Results	129
9.4.1	Batch 1: Admissions from January 2014	129
9.4.1.1	Women with incident ovarian cancer in cancer registry data	129
9.4.1.2	Women without incident ovarian cancer in cancer registry data	130
9.4.2	Batch 2: Admissions up to December 2013	132
9.4.3	Planned protocol for slide review and sample processing	133
9.5	Discussion	134

9.5.1	Challenges	135
9.5.2	Conclusion	136
10	Discussion	137
10.1	Overview	137
10.2	Aim of my DPhil work	137
10.3	Main findings from this thesis	138
10.3.1	Tubal ligation and risk of ovarian cancer histotypes	138
10.3.2	Tubal ligation and risk of other cancers	139
10.3.3	Reproductive factors and risk of ovarian cancer histotypes	140
10.3.4	Image analysis of the fallopian tubes	140
10.3.5	MWS Ovarian cancer tissue pilot study	141
10.4	Findings in context: the new hypotheses of ovarian cancer origins	142
10.5	Strengths and limitations of the study	145
10.6	Recommendations for future research	148
10.7	Conclusion	150
	References	152
	Appendix A Million Women Study recruitment questionnaire	173
	Appendix B Ethical approval for the MWS ovarian cancer tissue pilot study	178
B.1	MWS Appendix 28: Ovarian tumour tissue substudy	178
B.2	Letter confirming ethical approval for the MWS ovarian tumour tissue pilot study	187

CONTENTS

Appendix C Papers published from DPhil work	191
C.1 Tubal ligation and ovarian cancer risk in a large cohort: Substantial variation by histological type (International Journal of Cancer, 2016)	191
C.2 Tubal ligation and incidence of 26 site-specific cancers in the Million Women Study (British Journal of Cancer, 2016)	201
Appendix D Papers published during DPhil as part of collaborations with others	207
D.1 Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies (Lancet, 2015)	207
D.2 Past cervical intraepithelial neoplasia grade 3, obesity, and earlier menopause are associated with an increased risk of vulval cancer in postmenopausal women (British Journal of Cancer, 2016)	216
D.3 Premalignant SOX2 overexpression in the fallopian tubes of ovarian cancer patients: Discovery and validation studies (EBioMedicine, 2016)	225

LIST OF TABLES

4.1	Classification of the histological types of ovarian cancer	38
5.1	Characteristics of the study population at recruitment, and details of follow-up for ovarian cancer, according to tubal ligation status	53
6.1	Characteristics of the study population at recruitment, and details of follow-up, according to tubal ligation status	70
7.1	Characteristics of the study population at recruitment, and details of follow-up for ovarian cancer, according to parity	91
8.1	Table comparing statistics for % of nuclei showing 3+ staining in Phase I validation sample set (n=40), when the tubal epithelium is identified by hand or by macro.	110
8.2	Table comparing statistics for % of nuclei showing 3+ staining in Phase II validation sample set (n=17), when the stained tubal epithelial cells are identified and counted either manually by individuals (KG & DW), or with tubal epithelium identified by a macro or hand-annotation in Aperio, and the counting of stained cells performed by the ‘nuclear V9’ algorithm in Aperio	114
8.3	Table comparing statistics for absolute differences in the % of nuclei showing 3+ staining in validation sample set (n=17), comparing differences between two human observers, and the differences when the tubal epithelium is identified by a macro vs hand-annotation in Aperio	122

LIST OF TABLES

9.1	Batch 1: Table showing the results of comparing women with a diagnosis of ovarian cancer in electronic hospital records to data from cancer registries and pathology reports	132
-----	--	-----

LIST OF FIGURES

2.1	Histological images of ovarian cancer histotypes: A - Low-grade serous, ²⁹ B - High-grade serous, ³⁰ C - Endometrioid, ³¹ D - Clear cell, ³² E - Mucinous. ³³ (Source: Juan Rosai's collection of surgical pathology specimens http://www.rosaicollection.org/)	8
3.1	Map showing global age-standardised incidence rates of ovarian cancer (Source: GLOBOCAN 2012 ⁶)	14
3.2	Bar chart showing global age-standardised incidence and mortality rates of ovarian cancer (Source: GLOBOCAN 2012 ⁶)	14
3.3	Fixed-effects meta-analysis of the published prospective studies on talc use and ovarian cancer risk ^{134,135}	25
4.1	Flowchart showing detailed standard exclusions for the Million Women Study	37
5.1	Graph showing percentage of women reporting female sterilization (tubal ligation) by age and year (Source: Office for National Statistics ^{164–166}) . .	42
5.2	Participants inclusion/exclusion flow diagram for analyses of tubal ligation and ovarian cancer	49
5.3	Relative risk of ovarian cancer histotypes in women with versus without a history of tubal ligation	54

LIST OF FIGURES

5.4	Relative risk of ovarian cancer (overall and for the four main histotypes) in relation to the timing of tubal ligation	55
5.5	Relative risk of ovarian cancer amongst women with versus without tubal ligation, by subgroups	56
6.1	Participants inclusion/exclusion flow diagram for analyses of tubal ligation and all cancers	67
6.2	Relative risk of 26 site- or type-specific cancers in women with versus without a history of tubal ligation	71
6.3	Relative risk of lung cancer in women with versus without a history of tubal ligation, by smoking status	72
7.1	Participants inclusion/exclusion flow diagram for analyses of reproductive factors and risk of ovarian cancer	87
7.2	Relative risk of ovarian cancer by parity	92
7.3	Relative risk of histological types of ovarian cancer by parity	93
7.4	Relative risk of histological types of ovarian cancer by number of births and age at first and last births	94
7.5	Relative risk of ovarian cancer amongst parous vs nulliparous women, by subgroups	95
7.6	Relative risk of ovarian cancer with increasing duration of breastfeeding .	96
7.7	Relative risk of ovarian cancer histotypes per 12-months breastfeeding . .	97
8.1	Fallopian tube section stained for SOX2	111
8.2	Fallopian tube epithelium, as identified by the macro	111
8.3	Phase I (n=40): Scatter plot, showing the correlation between the percentage of epithelial cells scoring 3+ for staining for SOX2, when the epithelium was identified by hand-annotation versus by the macro	112

8.4	Phase I (n=40): Bland-Altman plot, showing the agreement between the percentage of epithelial cells scoring 3+ for staining for SOX2, when the epithelium was identified by hand-annotation versus by the macro	113
8.5	Phase II (n=17): Scatter plot showing the correlation between two observers (KG and DW) in counting total number of epithelial cells	114
8.6	Phase II: Scatter plot showing the correlation between two observers (KG and DW) in counting the percentage of epithelial cells stained for SOX2 by immunohistochemistry	115
8.7	Phase II: Bland-Altman plot, showing the agreement between the percentage of epithelial cells scoring 3+ for staining for SOX2, comparing a) manual counts vs b) the original ‘gold-standard’ method of identifying the epithelium by hand-annotation in Aperio, followed by automated cell-counting in Aperio	116
8.8	Phase II: Bland-Altman plot, showing the agreement between the percentage of epithelial cells scoring 3+ for staining for SOX2, comparing a) manual counts vs b) epithelial identification by the macro, followed by automated cell-counting in Aperio	116
8.9	Phase II: Bland-Altman plot, showing the agreement between the percentage of epithelial cells scoring 3+ for staining for SOX2, comparing a) identifying the epithelium by hand-annotation in Aperio, followed by automated cell-counting in Aperio vs b) epithelial identification by the macro, followed by automated cell-counting in Aperio	117
8.10	Fallopian tube section stained for SOX2, with the epithelium identified by hand-annotation, and staining assessed by Aperio software	118
8.11	Fallopian tube section stained for SOX2, with epithelium identified by the macro in FIJI/ ImageJ, and staining assessed by Aperio software	118
8.12	Fallopian tube section, stained for SOX2, with epithelium identified by hand-annotation, and staining assessed by Aperio software	120

LIST OF FIGURES

8.13 Fallopian tube section, stained for SOX2, with epithelium identified by macro in FIJI/ImageJ, and staining assessed by Aperio software	120
8.14 Fallopian tube section, stained for SOX2, with stained epithelial nuclei marked by manual counts in FIJI/ImageJ	121
8.15 Fallopian tube section, stained for SOX2, with unstained epithelial nuclei marked by manual counts in FIJI/ImageJ	121

LIST OF ABBREVIATIONS

95% CI		95% Confidence Interval
BMI		Body Mass Index
BRCA1/2		Tumour-suppressor genes, mutations in which confer increased susceptibility to early-onset breast cancer and ovarian cancer
CEU		Cancer Epidemiology Unit (at the University of Oxford)
FFPE		Formalin-fixed, paraffin-embedded (the standard method of tissue preparation for histopathology)
GWAS		Genome-Wide Association Studies
HES		Hospital Episode Statistics
HGSOC		High-Grade Serous Ovarian Cancer
HPV		Human Papilloma Virus
HR		Hazard Ratio (from a Cox regression model, often referred to simply as Relative Risk)
HRT or HT	. .	Hormone Replacement Therapy (also known as ‘menopausal hormones’)
HSCIC		Health and Social Care Information Centre
ICD-10		International Classification of Diseases, revision 10
ICD-O		International Classification of Diseases for Oncology
IQR		Interquartile Range
ISD Scotland	. .	Information Services Division Scotland

LIST OF ABBREVIATIONS

MWS		The Million Women Study
NCIN		National Cancer Intelligence Network
NDPH		Nuffield Department of Population Health
NHS		National Health Service
NHSCR		National Health Service Central Register
NOS		Not Otherwise Specified (used in histological classifications)
NSAIDS		Non-steroidal anti-inflammatory drugs
OCP		Oral Contraceptive Pill
ONS		Office for National Statistics
Para-1		Parity of one (i.e. a woman who has had one full-term pregnancy)
PHE		Public Health England
PLCO		Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial
RR		Relative Risk
SD		Standard Deviation
SMR		Scottish Morbidity Records
SOX2		SRY (sex determining region Y)-box 2, a gene encoding a transcription factor important in maintaining pluripotency of stem-cells
TVU		Trans-vaginal ultrasound
UK		United Kingdom
UKCTOCS	. . .	UK Collaborative Trial of Ovarian Cancer Screening
USA		United States of America
vs		versus

THE ROLE OF THE AUTHOR IN THIS THESIS

The work presented in this thesis is my own, conducted under the supervision of Professor Dame Valerie Beral and Professor Ahmed A. Ahmed, except where otherwise stated. The initial concept of investigating the association between lifestyle and reproductive factors and ovarian cancer by histological type in the Million Women Study was originally developed during discussions with Professor Beral. Throughout my DPhil, I have benefited greatly from advice from colleagues at the Cancer Epidemiology Unit (CEU, led by Prof Beral) and the Ovarian Cancer Cell Laboratory (led by Prof Ahmed). The specific contribution of individuals to different parts of my work is outlined below.

Chapter 5: Tubal ligation and risk of ovarian cancer histotypes

The idea to investigate tubal ligation as a risk factor for ovarian cancer was developed in concert with Professor Beral, in the light of existing literature.¹ The technique used for conducting tests for heterogeneity by histotype, using a competing risks approach² (also used in Chapter 7) was adapted from STATA code originally written by Kirstin Pirie (a statistical programmer at CEU). The information on tumour grade (used in this and subsequent chapters) was extracted from NCIN (National Cancer Intelligence Network) data by Carol Hermon, a data analyst at CEU.

Chapter 6: Tubal ligation and risk of 26 site-specific cancers

The idea to investigate the association between tubal ligation and risk of other cancer types was initially provoked by discussions at Professor Ahmed's weekly lab meeting - and in particular by questions from Dr Yiyan Zheng (a post-doctoral researcher). My approach to the analyses was influenced by previous work in the Million Women Study by Associate Professor Naomi Allen³ (a senior epidemiologist) and Professor Gillian Reeves (a statistical epidemiologist).⁴ I also benefited from discussions with Dr Kate Coffey (a former DPhil student at CEU).⁵

Chapter 7: Reproductive factors and risk of ovarian cancer histotypes

I had originally begun investigating this association as part of a summer project for an MSc course in 2011. However, as the data have now changed substantially (now with approximately double the number of ovarian cancers that were then available), and some of my analytical/ statistical methods have also altered, the work presented for this thesis is substantially different to that previously submitted for the MSc dissertation. The piecewise regression model used to model parity in this chapter was adapted from code originally written by Dr Isobel Barnes (a medical statistician).

Chapter 8: Image analysis of the fallopian tubes

The idea to develop and validate an algorithm for the identification of the fallopian tube epithelium was primarily my own, but developed in concert with Professor Ahmed. The macro itself was written by Dr Dominic Waithe (a computational biologist). The validation studies were performed primarily by myself, with Dr Waithe assisting in performing manual cell-counts (so as to have two independent individuals).

Chapter 9: MWS Ovarian Cancer Tissue Pilot Study

The idea to initiate a pilot study of obtaining archived tissue samples from MWS participants with ovarian cancer in Oxford was my own, but developed with input and discussion from many people, including Professor Beral, Professor Ahmed, Professor Reeves, Professor Jane Green (a clinical epidemiologist), and Associate Professor Ruth Travis (a senior molecular epidemiologist). The project also benefited from discussions with Dr Sunanda Dhar and Dr Sanjiv Manek, expert consultant gynaecological pathologists. The practicalities of ensuring appropriate ethical and research approvals were in place were greatly helped by Krys Baker (unit research coordinator at CEU), Lynden Guiver (study coordinator for the Million Women Study), and Professor Green. Identification of MWS participants who were treated for ovarian cancer in Oxford, and preparation of data for linkage, was performed by Carol Hermon. Linkage of the list of relevant MWS participants to the Oxford histopathology database was performed by the staff of the Oxford Centre for Histopathology Research (OCHRe), particularly Dr Stephanie Jones and Rebecca Davies. The development of the data transfer and data security policies/ mechanisms required for this process was the work of many people, including Dave Ewart (senior IT and information security manager at CEU) and Professor Green.

CHAPTER 1

THESIS INTRODUCTION

1.1 RATIONALE

Ovarian cancer is the seventh leading cause of cancer mortality amongst women world-wide, and the fourth leading cause amongst women in the UK.⁶ For UK women diagnosed with ovarian cancer, the age-standardised five-year relative survival is 31.0% (95% CI: 30.6-31.5).⁷ This relatively poor survival is probably attributable, at least in part, to relatively late stage at diagnosis: of those women with known stage, about 60% have advanced disease with metastatic spread (stage 3 or 4) at diagnosis.⁸ Although there have been several attempts to develop screening programmes for early detection of ovarian cancer (in the hope that this would improve survival), these have thus far not shown sufficient mortality benefit for immediate implementation at a population level.⁹⁻¹¹

Efforts at prevention have been limited by the relatively few established, easily-modifiable, risk-factors for ovarian cancer. There are a few high-penetrance genetic risk factors (e.g. mutations in the genes *BRCA1* or *BRCA2*, associated with familial breast and ovarian cancer syndrome¹²⁻¹⁴), for which prevention focuses on risk-reducing surgery.¹⁵ Genome-wide association studies (GWAS) have identified some additional sus-

ceptibility loci,¹⁶ but these have thus far not led to widely-accepted, clinically-useful, risk-prediction models.

Various lifestyle and reproductive factors have been shown or suggested to be associated with an increased risk of ovarian cancer. Some of these are potentially modifiable (e.g. body-mass index,¹⁷ or use of menopausal hormones¹⁸), but many others are related to decisions around the use of contraceptives,¹⁹ or other reproductive factors,²⁰ which occur in early adulthood and may thus be less easily modified.

An important development in our understanding of ovarian cancer has also been the increasing evidence that the different histological subtypes (histotypes) of ovarian cancer may have different origins, and show different molecular signatures.^{21–24} The vast majority of ovarian cancers are epithelial tumours, of which the four most common histotypes are serous, mucinous, endometrioid, and clear cell tumours.²⁵ For serous tumours, the most common histotype, it has been further suggested that they be divided into low-grade serous tumours (consisting of serous borderline tumours and low-grade serous carcinoma) and high-grade serous carcinoma, and that these represent two quite distinct pathways of tumour development.

There is thus a need to expand and develop epidemiological analyses of ovarian cancer, taking account of the probable importance of potential variation by histological type. This would not only aid in understanding the risk factors better, with a view to possible modification and prevention, but might also aid in elucidating the underlying biology and development of the different histotypes. Although several individual studies and large consortia have attempted this, their findings have often been limited by either retrospective designs (with consequent risk of bias)^{1,26} or insufficient numbers of cases of the less-common histotypes.^{20,27}

1.2 RESEARCH AIMS

The research documented in this thesis is primarily concerned with investigating reproductive and lifestyle risk factors for ovarian cancer, exploring potential variation in these associations by histological tumour type, and interpreting the results in the light of new hypotheses of the origins of ovarian cancer.

The thesis also outlines current and ongoing research that aims to develop and establish resources and methods for exploring the risk factors and origins of ovarian cancer using tissue samples.

1.3 OUTLINE OF THE THESIS

Chapter 1 provides an overview of the thesis. It outlines the rationale for the thesis – namely increasing evidence that ovarian cancer is a heterogeneous disease, but incomplete knowledge of the variation of risk factors between the histological subtypes (as detailed in Chapters 2 and 3). It also outlines the approach taken to investigate these issues – primarily epidemiological analyses of the risk factors for ovarian cancer, split by histological subtype (Chapters 5 to 7), but also some more exploratory tissue-based work (Chapters 8 and 9).

Chapter 2 reviews the increasing evidence that ovarian cancer is a heterogeneous disease, with histological subtypes (most commonly, serous, endometrioid, clear cell and mucinous) that vary not only in microscopic appearance, but also in genetic/ molecular characteristics, and possibly in cell of origin. In particular, the chapter briefly reviews the evidence for the hypotheses that many high-grade serous ovarian cancers may originate from precursor lesions in the fallopian tube epithelium, while some endometrioid and clear cell tumours may arise from endometriosis.

Chapter 3 reviews the known epidemiology of ovarian cancer, and the increasing evidence that the histological subtypes of ovarian cancer have different population-level risk factors. This includes findings from the Million Women Study (one of the largest prospective cohort studies in the world, and the main source of data for this thesis), and the Collaborative Group on Epidemiological Studies of Ovarian Cancer. It also highlights the many important factors that have previously been suggested by other studies to be associated with ovarian cancer risk, but which have thus far not been explored by histotype in the Million Women Study – including tubal ligation, parity, and breastfeeding. Some of these exposures are particularly relevant for the new hypotheses of tubal or endometriotic origins for certain histotypes. This explains the need for the present study.

Chapter 4 details the methods of the Million Women Study (MWS), a large prospective cohort of UK women, which is the main data source used for the epidemiological analyses in this thesis. The chapter also outlines the main methods used for the epidemiological analyses – namely, Cox regression (survival time) analyses, adjusted for potential confounding factors.

Chapter 5 investigates the association between tubal ligation (female sterilization) and the risk of ovarian cancer, by histological tumour type. This study is a prospective analysis in the Million Women Study cohort, of 8,031 cases of ovarian cancer that had accrued by 31st December 2013, amongst women without previous cancer or bilateral oophorectomy. The work outlined in this chapter was published in the *International Journal of Cancer* (March 2016, Vol. 138, Issue 5, pages 1076-1084), and is also reproduced in Appendix C.1.

Chapter 6 investigates the association between tubal ligation (female sterilization) and the risk of 26 different cancer types, including cancers of the fallopian tube and peritoneum (which have been hypothesised to have similar origins to those of the most

common form of ovarian cancer). This study is a prospective analysis in the Million Women Study cohort, of cancer cases that had accrued by 31st December 2013, amongst women without previous cancer. The work outlined in this chapter was published in the British Journal of Cancer (May 2016, Vol. 114, Issue 9, pages 1033-1037), and is also reproduced in Appendix C.2.

Chapter 7 investigates the association between reproductive factors (parity and breastfeeding) and risk of ovarian cancer, by histological tumour type. This study is a prospective analysis in the Million Women Study cohort, of 8,719 cases of ovarian cancer that had accrued by 31st December 2014, amongst women without previous cancer or bilateral oophorectomy. A manuscript relating to this work has been prepared for publication.

Chapter 8 briefly introduces a collaborative project of work in which I am involved, that is investigating changes in protein expression (as assessed by immunohistochemistry) in the fallopian tube epithelial cells of women with the high-grade serous histotype of ovarian cancer (which is hypothesised to arise from the fallopian tubes), compared to women with other benign or malignant conditions. The chapter then describes a sub-project, of developing and validating an algorithm ('macro') for use in image analysis software, which will automate the identification of the fallopian tube epithelial layer, so as to facilitate analysis of scanned slides.

Chapter 9 describes a pilot study of identifying, retrieving, obtaining and processing pathology reports and tissue samples from Million Women Study participants who were diagnosed and treated for ovarian cancer in the Oxford region. I initiated the pilot study during my DPhil, and at the time of writing it is still in progress.

1.4 PUBLICATIONS RELATING TO THIS THESIS

At the time of writing, two manuscripts have been published from work described in this thesis. These are "Tubal ligation and ovarian cancer risk in a large cohort: Substantial variation by histological type", and "Tubal ligation and incidence of 26 site-specific cancers in the Million Women Study" (see Appendix C). My co-authors for the papers provided advice on the content and methodology, and read the papers both prior to submission and at revision - and so to this extent the papers may not be considered to be entirely my own work. A manuscript has also been prepared for publication related to the work outlined in Chapter 7.

Other papers to which I have contributed during the course of my DPhil, but which are not primarily based on material presented as part of this thesis, are included in Appendix D.

CHAPTER 2

OVARIAN CANCER PATHOLOGY

2.1 INTRODUCTION

As with the majority of cancers, ovarian cancer has historically been divided into different types, largely on the basis of microscopic morphological appearance. However, there is increasing evidence that the different histological subtypes (histotypes) of ovarian cancer may have different origins, and show different molecular signatures.^{21–24}

The vast majority of ovarian cancers are epithelial tumours (carcinomas), of which the four most common histotypes are serous, mucinous, endometrioid, and clear cell tumours.²⁵ In a European-wide study of ovarian cancer based on data from cancer registries, the largest specific histotype was serous tumours (accounting for 34% of ovarian cancer cases), followed by endometrioid (9%), mucinous (8%), and clear cell (3%); stromal and germ cell tumours, sarcomas, and squamous cell tumours together accounted for <5%; the remainder were classified as undifferentiated (25%), or of ‘other’ or ‘Not Otherwise Specified’ types.⁷

Serous ovarian carcinomas morphologically resemble the epithelium lining the fal-

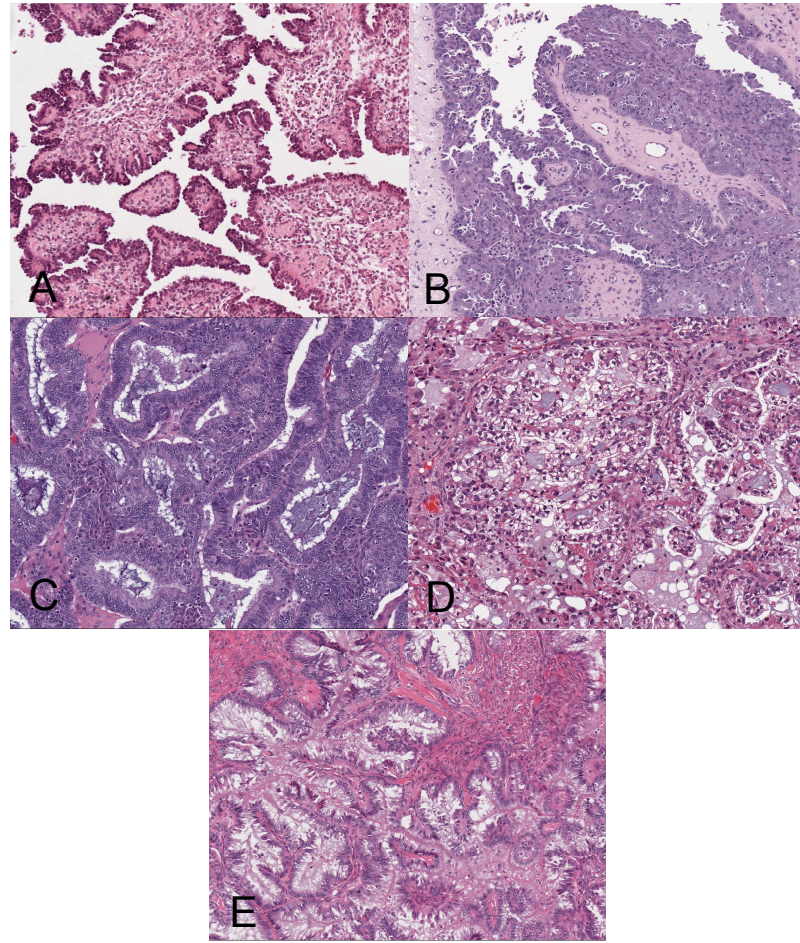


Figure 2.1: Histological images of ovarian cancer histotypes: A - Low-grade serous,²⁹ B - High-grade serous,³⁰ C - Endometrioid,³¹ D - Clear cell,³² E - Mucinous.³³ (Source: Juan Rosai's collection of surgical pathology specimens <http://www.rosaicollection.org/>)

lopian tube; low-grade (Figure 2.1 A) and high-grade serous carcinomas (Figure 2.1 B) are distinguished largely based on nuclear atypia and mitotic activity.²⁵ Endometrioid ovarian carcinomas (Figure 2.1 C) resemble the endometrium (the lining of the uterus). Clear cell carcinomas (Figure 2.1 D) tend to have clear-staining cytoplasm, attributed to abundant glycogen.²⁸ Mucinous carcinomas (Figure 2.1 E) tend to resemble either the lining of the gastrointestinal tract, or the glandular part of the uterine cervix.

2.2 NEW HYPOTHESES REGARDING THE ORIGINS OF SEROUS TUMOURS

Serous tumours have traditionally been divided into serous borderline tumours and invasive serous carcinomas, and serous carcinomas have customarily been graded using a three-tier system (low, medium or high-grade). As with several other cancer types, it was presumed that, in most cases, high-grade serous ovarian carcinomas had developed from low-grade serous ovarian carcinomas. However, it is now hypothesised that high-grade serous ovarian carcinomas are quite different entities, and in most cases will have developed along an entirely different pathway to serous borderline tumours and low-grade serous carcinomas (which possibly are different steps along the same pathway).²²

High-grade serous ovarian cancer is hypothesised to arise from precursor lesions within the fallopian tubal epithelium, particularly in the fimbrial end adjacent to the ovary, and subsequently seed to the ovary.²² Much of the initial evidence for this hypothesis came from studies showing putative precursor lesions within the fallopian tubes (such as focal areas of epithelium featuring mutations in the tumour-suppressor gene *TP53* or aberrant expression of its protein p53, and/or areas of dysplasia or serous intraepithelial carcinoma) in prophylactic salpingoophorectomy specimens from women with an increased genetic risk of ovarian cancer, due to inherited *BRCA1/2* mutations.³⁴⁻³⁷

Serous tubal intraepithelial carcinoma (STIC), the putative precursor lesion, has also been noted in the fallopian tubes of about 20-60% of women with high-grade serous ovarian cancer.³⁸⁻⁴⁷ Some of these studies also found similar lesions in women with apparent fallopian tubal carcinoma or primary peritoneal carcinoma, and so it has been hypothesised that STIC lesions may act as precursors to high-grade serous carcinomas found in the fallopian tubes, ovary, and peritoneum.⁴⁰ Clinical guidelines for staging and treating ovarian cancer reflect this, now advising that high-grade serous cancers of the ovary, fallopian tube and peritoneum should be considered collectively.^{48,49}

STICs have also been identified in a small proportion of salpingoophorectomy specimens from women not known to have ovarian cancer or to be at increased genetic risk, in some cases with accompanying incidental invasive high-grade serous carcinoma.^{50–52} The putative precursor lesions in many cases show molecular changes that are similar or identical to those seen in co-existing high-grade serous ovarian cancer, particularly in terms of *TP53* mutations³⁸ (mutations in the gene for p53 being characteristic and ubiquitous in high-grade serous ovarian cancer⁵³), which argues against STICs merely being part of a general pro-cancerous ‘field change’.

The origins of low-grade serous tumours are considerably more uncertain. In many cases, there appears to be a step-wise progression from ovarian epithelial inclusion cysts, through benign cystadenomas and serous borderline tumours, to low-grade serous carcinoma. Evidence for this includes the frequent co-existence of serous borderline tumours with low-grade serous carcinomas,⁵⁴ and similar molecular changes: both serous borderline tumours and low-grade serous carcinomas commonly feature mutations in *BRAF* or *KRAS*,⁵⁵ but rarely in *TP53*,⁵⁶ in contrast to high-grade serous carcinoma. It has been suggested that some low-grade serous tumours may also originate indirectly from tubal epithelial cells (perhaps from benign or hyperplastic tubal epithelial cells which become trapped in the ovary as epithelial inclusion cysts at the time of ovulation, possibly at a young age),^{57,58} but the evidence for this is not well established.

2.3 NEW HYPOTHESES REGARDING THE ORIGINS OF ENDOMETRIOID AND CLEAR CELL OVARIAN CANCER

Some endometrioid and clear cell ovarian cancers are thought to develop from endometriosis,^{22,23,59} deposits of endometrial-type glands and stroma that are found outside of the uterus, most commonly on the surface of the ovaries, fallopian tubes, and other serosal surfaces of the pelvic organs. This hypothesis is based primarily on histopathologi-

cal studies, which show that endometriosis is often seen near to these tumour types, and that there is often an observable transition between endometriosis and clear cell or endometrioid ovarian cancer (and more frequently than with other types of ovarian cancer).^{60–66} Estimates vary for the prevalence of endometriosis amongst women with different ovarian cancer histotypes. Studies among women in Europe and the USA suggest that endometriosis is seen nearby in about 10-50% of endometrioid and clear cell tumours, but in less than 5% of serous or mucinous tumours; prevalence estimates for endometriosis in ovarian cancer tend to be higher in Japan.^{64,67}

In addition, where endometrioid or clear cell ovarian cancers have arisen within or adjacent to endometriosis, common genetic lesions (particularly loss of heterozygosity events) are often detected within both the tumours and the endometriosis, consistent with a common lineage.^{68–70} Somatic mutations in the gene *ARID1A* (which may function as a tumour-suppressor gene) have been identified in 30-80% of endometrioid and clear cell tumours, and were also present in contiguous endometriosis in some patients, but were not found in high-grade serous ovarian carcinomas.^{71–74} Loss of expression of the *ARID1A*-encoded protein BAF250a is also seen in some cases of endometrioid and clear cell ovarian cancer, and contiguous endometriosis.⁷⁵ These findings would be consistent with the hypothesised continuum between endometriosis and endometrioid and clear cell ovarian cancers.

Other molecular changes that have been reported to be associated with endometrioid ovarian carcinoma include microsatellite instability (10-20% of cases), and mutations of the genes *PTEN* (10-20%) and *CTNNB1* (beta-catenin, up to 40-50%).^{74,76} Mutations in *PIK3CA* have been reported in about 30-50% of clear cell and endometrioid ovarian carcinomas,^{71,74,77} and may co-exist with *ARID1A* mutations.⁷⁸

2.4 HYPOTHESES REGARDING THE ORIGINS OF MUCINOUS TUMOURS

The origins of mucinous tumours are considerably less certain. It has been suggested that, particularly in older case series, many cases of mucinous ‘ovarian’ cancer may have been misdiagnosed metastases from the gastrointestinal tract.⁷⁹ Others have suggested that mucinous ovarian cancers may share a common origin with Brenner tumours (transitional cell tumours, which resemble urothelium lining the urinary tract), and may arise from Walthard cell nests (benign microscopic nests of transitional cells, found at the tubal-peritoneal junction).^{80,81} Some gastrointestinal-type mucinous tumours can also develop from mature ovarian cystic teratomas.⁸²

2.5 IMPLICATIONS FOR THIS THESIS

Given these new hypotheses of the potentially disparate origins of the ovarian cancer histotypes, it is important for epidemiological studies of ovarian cancer that the different histotypes should be examined separately - and so this thesis will look particularly at possible heterogeneity by histotype.

CHAPTER 3

OVARIAN CANCER EPIDEMIOLOGY

3.1 OVARIAN CANCER INCIDENCE AND MORTALITY

Ovarian cancer incidence rates show considerable global variation, with higher incidence seen particularly within North America, Europe, and Australasia, and lower incidence in sub-Saharan Africa and Asia (see Figures 3.1 and 3.2).

Within the UK, ovarian cancer is the sixth most common incident cancer amongst women (excluding non-melanoma skin cancer), with an age-standardised incidence rate of 11.7 per 100,000, and the fourth most common cause of cancer mortality (age-standardised rate of 5.7 per 100,000).⁶ The estimated cumulative lifetime risk of ovarian cancer (up to age 75) for women in the UK is 1.3%.⁶

For women in the UK diagnosed with ovarian cancer, the age-standardised 5-year relative survival has been estimated to be 31.0% (95% CI: 30.6-31.5).⁷ This relatively poor survival is partly attributable to an average late stage at diagnosis: about 60% of UK women with ovarian cancer have advanced disease with metastatic spread (stage 3 or 4) at diagnosis.⁸

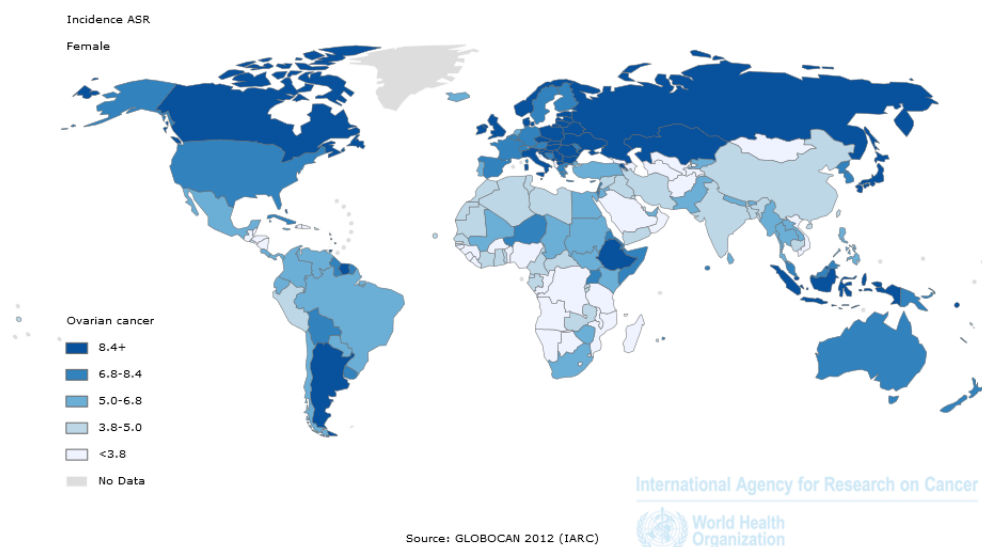


Figure 3.1: Map showing global age-standardised incidence rates of ovarian cancer (Source: GLOBOCAN 2012⁶)

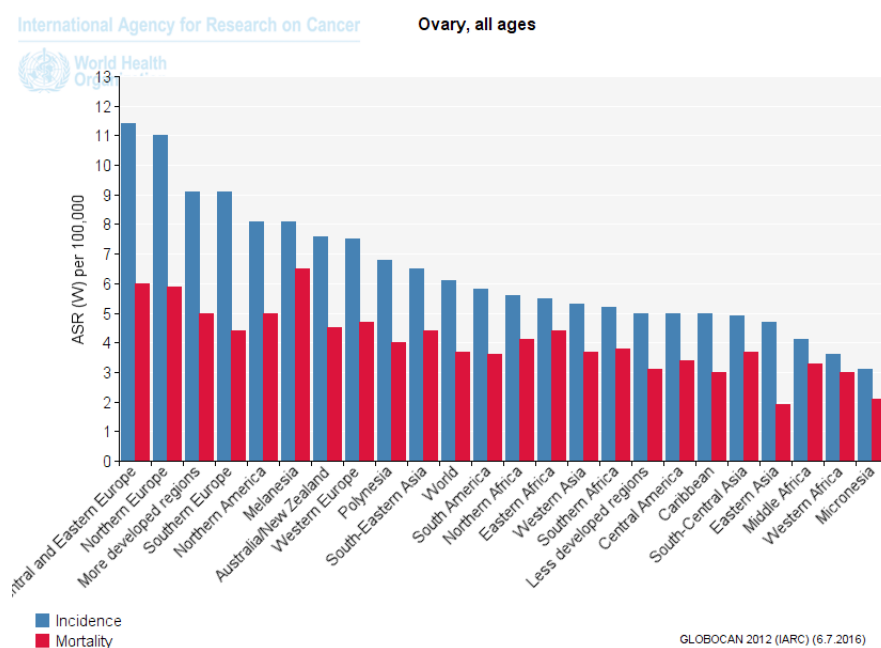


Figure 3.2: Bar chart showing global age-standardised incidence and mortality rates of ovarian cancer (Source: GLOBOCAN 2012⁶)

3.2 RANDOMISED TRIALS OF SCREENING

In an effort to diagnose ovarian cancer at an earlier stage, when there is a greater potential for treatment to be curative, there have been several randomised controlled trials of screening for ovarian cancer, generally using either trans-vaginal ultrasound and/or a blood test for the tumour marker CA-125. The two largest trials published thus far are the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) and the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS).

3.2.1 PLCO

In the ovarian cancer arm of the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO), of 78,216 women in the USA aged 55-74 who were randomized, 39,111 were allocated to the control group to receive usual care (4,807 were not included in analyses as they had previously had bilateral oophorectomy), and 39,105 were allocated to receive screening for ovarian cancer with annual trans-vaginal ultrasound (TVU) and CA-125 blood test for four years, followed by the blood test alone for two years.^{9,10}

After 12.4 years of median follow-up, as expected, there appeared to be a higher incidence of ovarian cancer cases diagnosed in the screening arm: 212 ovarian cancer cases (5.7 per 10,000 person-years) in the screening group, compared to 176 cases (4.7 per 10,000 person-years) in the control group, although this difference was not statistically significant (RR: 1.21, 95% CI: 0.99-1.48).¹⁰ However, there was no apparent stage shift: stage III and IV cancers accounted for the majority of cases in both the screening (77%) and control group (78%). There was also no evidence of a benefit in ovarian-cancer-specific mortality: after 118 deaths from ovarian cancer in the screening arm (3.1 per 10,000 person-years), and 100 deaths in the control arm (2.6 per 10,000 person-years), the adjusted relative risk for mortality was 1.18 (95% CI: 0.82-11.71).¹⁰

A later analysis suggested that there would still have been no significant mortality benefit, even if the CA-125 blood test had used the risk of ovarian cancer algorithm (ROCA, a statistical tool that determines ovarian cancer risk based on past and current CA-125 values), rather than an absolute cut-off of ≥ 35 U/mL.⁸³

3.2.2 UKCTOCS

In the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), 202,638 women aged 50-74 were randomised to one of three arms: 50% (101,359) to no screening, 25% (50,640) to screening with TVU and CA-125 blood-test (making use of the ROCA algorithm, rather than an absolute cut-off value), and 25% (50,639) to screening with TVU alone.¹¹ After a median follow-up of 11.1 years, ovarian cancer had been diagnosed in 1,282 women, of whom 649 had died of ovarian cancer.

There was no evidence of a significant benefit in the primary outcome of ovarian-cancer-specific mortality: screening with ultrasound alone had an 11% reduction in mortality compared to usual care (95% CI: -7 to 27, $p=0.21$), while combined screening with ultrasound and CA-125 blood test had a 15% reduction in mortality (95% CI: -3 to 30, $p=0.10$). The authors noted a significant mortality reduction in the combined screening group when they did an analysis excluding prevalent cases (i.e. those diagnosed at the first round of screening), but this was not pre-specified as the primary analysis.

3.3 GENETIC RISK

3.3.1 BRCA

It has long been known that some families carry an hereditary risk for ovarian cancer - most famously in the form of the hereditary breast and ovarian cancer syndromes due to

mutations in the tumour suppressor genes *BRCA1* and *BRCA2*, which were first mapped in the early 1990s.^{12–14} It is estimated that carriers of mutations in *BRCA1* have about a 45% cumulative risk of developing ovarian cancer by age 80 (and about a 67% cumulative risk of breast cancer), while those with *BRCA2* mutations have an estimated 12% cumulative risk of ovarian cancer by age 80 (and 66% risk of breast cancer).^{15,84,85} The majority of ovarian cancers that develop amongst women with *BRCA1/2* mutations are of the high-grade serous histotype.⁸⁶ Standard clinical guidelines recommend that women with known deleterious mutations of *BRCA1* or *BRCA2* consider prophylactic risk-reducing surgery in the form of bilateral salpingoophorectomy (removal of both ovaries and fallopian tubes) and bilateral mastectomy (removal of both breasts), after a certain age (e.g. 35 to 45), and when they have completed childbearing.^{87,88}

3.3.2 Other genetic risks

Other inherited cancer syndromes with an increased risk of ovarian cancer include Lynch syndrome (characterised by germline mutations in one or more of the DNA mismatch repair genes *MLH1*, *MSH2*, *MSH6* and *PMS2*, leading to increased risks of colorectal, endometrial and ovarian cancers),⁸⁹ and Li-Fraumeni syndrome (characterised by increased risk of cancer at multiple sites, including the ovary, due to germline mutations in the tumour-suppressor gene *TP53*⁹⁰). Other germline mutations that have been associated with risk of ovarian cancer include *BARD1*, *BRIP1*, *CHEK2*, *MRE11A*, *NBN*, *PALB2*, *RAD50* and *RAD51C*, many of which are involved in the Fanconi anaemia pathway.⁹¹

Genome-wide association studies (GWAS) have identified about 22 further lower-penetrance genetic susceptibility loci that appear to be associated with ovarian cancer risk.^{16,92–100} Suggestive associations have recently been reported at further loci, though these did not reach traditional statistical significance for GWAS studies.^{101,102}

3.4 FINDINGS FROM THE COLLABORATIVE GROUP ON EPIDEMIOLOGICAL STUDIES OF OVARIAN CANCER

Probably the most reliable and comprehensive source of information on the risk factors for ovarian cancer comes from the Collaborative Group on Epidemiological Studies of Ovarian Cancer ('the Collaboration'). This is an international collaborative effort to conduct meta-analyses using individual-patient-level data from all epidemiological studies (whether published or not) that have information on ovarian cancer and certain potential risk factors. The Collaboration now has information on 31,000 women with ovarian cancer (plus a greater number of controls) from 58 studies worldwide, and is co-ordinated from the Cancer Epidemiology Unit at the University of Oxford. More information is available at the website <https://www.ceu.ox.ac.uk/research/epidemiological-studies-of-ovarian-cancer>.

I have been involved in the Collaboration both before and during my DPhil, and am a co-author (i.e. member of the writing committee) on two publications.^{18,103}

The Collaboration has thus far reported on associations between ovarian cancer risk and use of the oral contraceptive pill, body size, smoking, and use of menopausal hormones. A substantial reduction in risk of ovarian cancer has been shown amongst women taking the oral contraceptive pill, with an estimated 20% reduction in relative risk per 5 years of use (95% CI: 18-23%).¹⁹ There is an increased risk of ovarian cancer amongst women with a larger body size - not only for higher BMI, with an estimated 10% increase in relative risk per 5 kg/m² increase in BMI (RR: 1.10, 95% CI: 1.07-1.13, $p < 0.001$), but also with greater height, with a 7% increase in risk per 5 cm increase in height (RR: 1.07, 95% CI: 1.05-1.09, $p < 0.001$).¹⁷ Current or recent users of menopausal hormones have been reported to have nearly a 40% increased risk of ovarian cancer (RR: 1.37, 95% CI: 1.29-1.46), although this declines with greater time since use ceased.¹⁸

For smoking, the association is more complex: although current smokers had a small increased risk of ovarian cancer overall (RR: 1.06, 95% CI: 1.01-1.11, $p=0.01$), this was almost entirely driven by a large increase in the risk of mucinous tumours.¹⁰³

Although all of these associations had previously been reported on by at least some of the individual studies, the systematic synthesis of all the available evidence means that the findings of the Collaboration are more reliable and informative than any of the individual studies in isolation. An effort was made by the Collaboration to include all relevant studies, especially for analyses of the oral contraceptive pill and menopausal hormones. Including the totality of evidence, in so far as is possible, should help to avoid potential biases associated with emphasis just on certain studies. The large size of the Collaboration has also enabled it to explore potential heterogeneity in overall epidemiological associations, not only by timing or amount of exposure, but also by methodological factors (e.g. whether a study was retrospective or prospective) and by biological factors (e.g. by histological tumour type). These findings are discussed in more detail below.

3.4.1 Heterogeneity by histological type

As previously discussed in Chapter 2, there is increasing evidence that the different histological types of ovarian cancer, which were originally defined by microscopic morphology, also have different characteristic mutations, and possibly different origins. There has thus been great interest in exploring whether population-level risk factors also vary between the histotypes, and in a manner consistent with the new hypotheses of their disparate origins. The Collaboration, with its large number of cases of even relatively rare histotypes, is well-placed to investigate this, and has found several instances of heterogeneity by histotype.

For the oral contraceptive pill, there was about a 20% reduction in risk of ovarian

cancer per 5 years of use for serous, clear cell and endometrioid tumours, but there was no evidence of an association with risk of mucinous tumours; heterogeneity: $p=0.0007$.¹⁹

For height, there was about a 7% increase in risk per 5 cm increase in height for serous, clear cell, and mucinous tumours; although the estimate for endometrioid tumours did not show a significant increase in risk, tests for heterogeneity did not show a significant difference in the association by histotype.¹⁷ For BMI, each 5 kg/m² increase in BMI was associated with about a 10% increase in risk of mucinous and endometrioid tumours, but no significant association with clear cell tumours; for serous tumours, there was a significant difference in the association between BMI and invasive serous carcinoma (null) vs borderline serous tumours (about a 30% increase in risk per 5-unit increase in BMI); heterogeneity: $p<0.001$.¹⁷

For smoking, the associations by histotype are particularly interesting, as they vary in different directions. Current smokers appear to have no change in risk of serous tumours, and about a 20% reduction in risk of endometrioid and clear cell tumours, compared to never smokers; however, they have about a 50% increase in risk of invasive mucinous carcinoma, and over a doubling of risk of mucinous borderline tumours (heterogeneity: $p<0.0001$).¹⁰³

Menopausal hormone therapy appears to be associated with a substantial and significant increased risk of serous and endometrioid tumours (about a 40-50% increased risk amongst current or recent users compared to never-users), but either no association or a possible reduced risk of mucinous and clear cell tumours (heterogeneity: $p<0.0001$).¹⁸

Although the putative different origins of the different subtypes gives a reasonable theoretical framework for thinking that they might have different risk factors, the underlying biological explanations for the specific differences observed (e.g. a pronounced increase in risk of borderline serous tumours with increased BMI, a null association between mucinous tumours and the OCP, or an increased risk of mucinous tumours but

not others with smoking) is largely unclear.^{17,19,103}

3.4.2 Retrospective vs prospective studies

Another finding that has been important from the Collaboration is the potential for very different associations to be seen, depending on whether a study used a prospective versus a retrospective design.

In a prospective study, exposure history is ascertained prior to disease diagnosis - usually as part of a cohort design, in which healthy women are recruited, assessed in person or via questionnaire, and then followed up for incident disease or death. By contrast, in a retrospective study, exposure history is ascertained amongst women with a diagnosed disease (cases) and those without the disease (controls). Retrospective studies are therefore susceptible to additional biases, including selection bias (particularly as regards the difficult matter of how to select appropriate controls, including the choice of whether to recruit from the same hospitals as the cases, or from the general population) and recall bias (as a recent diagnosis of cancer might subconsciously alter a person's recollection of previous life experience). In addition, the poor survival for ovarian cancer means that rapidly-fatal cases may not be represented in retrospective studies, as there is often some time between diagnosis of cancer and interview for the study.

Given this potential for heterogeneity in the study types, the Collaboration has reported results separately for prospective and retrospective studies, and subdivided the latter by studies that used population controls versus hospital controls.

For some exposures, there appears to be relatively little heterogeneity by study type: for example, both prospective and retrospective studies find about a 20-30% reduction in risk of ovarian cancer amongst ever vs never users of the oral contraceptive pill.¹⁹

However, for other exposures, there were substantial differences by study design.

For example, for body size, retrospective studies with hospital controls found that increased height and BMI were associated with a *reduced* risk of ovarian cancer - in the opposite direction to the associations found in either retrospective studies with population controls, or prospective studies.¹⁷ The retrospective studies with hospital controls were thus excluded from subsequent analyses of ovarian cancer and body size, due to concerns that some of the controls might have been women with other diseases related to height or weight.¹⁷

The same issue was seen in analyses of smoking and ovarian cancer risk: retrospective studies with hospital controls found that smoking was associated with a substantially *reduced* risk of ovarian cancer (about a 20% reduction in risk amongst ever vs never smokers), whereas both prospective studies, and retrospective studies with population controls, found that smoking was associated with a small (6-8%) *increased* risk of ovarian cancer overall.¹⁰³ Again, the retrospective studies with hospital controls had to be excluded from future analyses, as it seemed likely that some of the hospital controls may have had smoking-related diseases¹⁰³ (which is perhaps unsurprising, given the wide range of diseases with which smoking is associated¹⁰⁴).

Retrospective studies also appeared to give rise to different results in analyses of menopausal hormone use and ovarian cancer risk. Although there was strong evidence of a substantial increased risk of ovarian cancer associated with use of menopausal hormones in prospective studies from both Europe and North America (about a 35% increased risk amongst current or recent users vs never-users), the association was null amongst the majority of retrospective studies (there was some evidence of an increased risk in the European retrospective studies, but not in those from North America or elsewhere).¹⁸ This difference was partly attributed to difficulties in selection of controls in many of the retrospective studies, but also to the relatively shorter durations of use of menopausal hormones in many of the retrospective studies.¹⁸ It is also possible that the cases in the

retrospective studies were not representative of ovarian cancer overall - for example, if a long time interval between diagnosis and recruitment/ interview for the study meant that many cases had died. This is a particular challenge with retrospective studies of ovarian cancer, given the relatively poor survival.

3.5 ASSOCIATIONS NOT YET REPORTED ON BY THE COLLABORATION

In addition to the associations reported on by the Collaboration (described above), there are several other factors which have been previously investigated as potential risk factors for ovarian cancer.

For some exposures, the Million Women Study has previously reported on associations with several different cancer types, including ovarian cancer. For example, analyses from MWS have shown no trend in risk of ovarian cancer with moderate alcohol intake,³ consumption of organic food,¹⁰⁵ or birth weight.¹⁰⁶

Many other studies have shown that ovarian cancer risk is reduced amongst women who have children,^{20,107–122} and some have suggested additional reductions in risk with breastfeeding, though the evidence is mixed^{123–126} (reviewed in more detail in Chapter 7).

Meta-analyses of retrospective studies have reported a reduced risk of ovarian cancer amongst women who have undergone tubal ligation (female sterilization)^{1,127} (discussed in more detail in Chapter 5).

There have been several observational epidemiological studies, which have reported a strong association between endometriosis and endometrioid and clear cell ovarian cancers. A pooled analysis of 21 prospective studies (by Wentzensen *et al.*) from the Ovarian Cancer Cohort Consortium included six studies with information on endometriosis, with 1,503 cases of ovarian cancer.²⁷ This analysis found heterogeneity between endometriosis

and the diagnosis of certain ovarian cancer histotypes, reporting that endometriosis was associated with a significantly increased risk of endometrioid, clear cell, and low-grade serous ovarian cancer, but no apparent increase in risk for mucinous or high-grade serous carcinomas.²⁷ Similar results were also reported by Pearce *et al.* of the Ovarian Cancer Association Consortium, who pooled data from 13 retrospective studies of ovarian cancer.¹²⁸

Evidence of an association with hysterectomy is varied, and there appears to be some change in the associations reported over time: studies in which the average year of diagnosis was prior to 2000 have tended to report a reduced risk of ovarian cancer with hysterectomy, while studies with more recent diagnoses have tended to report an increased risk.¹²⁹ Prospective studies have been reported to show more evidence of a reduction in risk of ovarian cancer following hysterectomy than retrospective studies; however, this may be influenced by the aforementioned secular trends, as the bulk of the available evidence from prospective studies was from pre-2000.¹²⁹

Some reports have suggested an increased risk of ovarian cancer amongst women who have used certain drugs to treat infertility (particularly those that cause ovarian stimulation/ hyper-ovulation). However, a systematic review, which considered 11 retrospective studies and 14 prospective studies, did not find convincing evidence of an increased risk of ovarian cancer amongst women treated for infertility.¹³⁰

Potential associations between diet and ovarian cancer risk have been explored by many studies, but there are no exposures for which there is consistent evidence of a substantial association with ovarian cancer (other than the aforementioned increased risk with body size),¹²⁶ though some have suggested an increased risk associated with higher fat consumption.^{131,132}

There have been several studies showing an association between use of talc (talcum powder) in the genital/ perineal area and risk of ovarian cancer, with a pooled analysis

of eight retrospective studies estimating about a 20% increase in the relative risk of ovarian cancer amongst users vs non-users (OR: 1.24, 95% CI: 1.15-1.33).¹³³ However, the majority of evidence in favour of this association has come from retrospective studies - which, as previously demonstrated by the Collaboration, may sometimes give incorrect results due to bias. Only two prospective studies have reported on the issue (the Nurses' Health Study¹³⁴ and the Women's Health Initiative Observational Study¹³⁵), and both found no significant association between perineal talc use and overall risk of ovarian cancer (combined RR: 1.07, 95% CI: 0.92-1.24, see Figure 3.3), and also found no evidence of a dose-response effect with duration of use. However, as each of the prospective studies only had fewer than 500 cases,^{134,135} compared to the 8,500 cases in the pooled analysis of eight retrospective studies,¹³³ they may be under-powered to detect a true association.

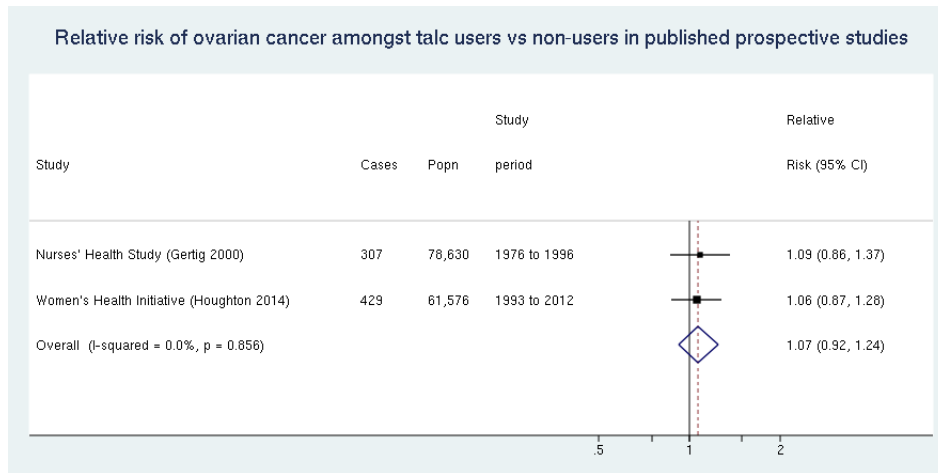


Figure 3.3: *Fixed-effects meta-analysis of the published prospective studies on talc use and ovarian cancer risk*^{134,135}

The association between talc use and ovarian cancer risk cannot yet be investigated in the Million Women Study, as a question on talc use was only included in a later resurvey questionnaire (about 12 years after recruitment). Thus, there has not yet been enough follow-up time for sufficient cases of ovarian cancer to accrue. However, this would be an important question to investigate in MWS in a few years' time.

Some studies have suggested an inverse association between age at menarche and ovarian cancer, with a recent meta-analysis finding about a 15% reduction in risk for the oldest versus the youngest age categories; this was just statistically significant in retrospective studies, but not in prospective studies.¹³⁶ Later age at menopause has been found by several studies to be associated with an increased risk of ovarian cancer: a pooled analysis from the Ovarian Cancer Cohort Consortium (including 19 studies with information on age at menopause) found an estimated 6% increase in risk per 5-year increase in age at menopause for all invasive ovarian cancer, with significant heterogeneity by histotype (an estimated 5% increase for serous tumours, and about a 20-40% increase for endometrioid and clear cell tumours, but no association for mucinous tumours).²⁷

Several studies have reported that there is a reduced risk of ovarian cancer amongst women who use aspirin and other non-steroidal anti-inflammatory drugs (NSAIDs, e.g. ibuprofen),^{137,138} though other studies have not found a significant association.¹³⁹

3.6 MECHANISMS

The biological mechanisms underlying the reported epidemiological associations are largely uncertain. One popular hypothesis is that ovarian cancer risk is at least partially driven by either hormones (e.g. oestrogens) or ovulation (whether through the hormonal effects of ovulation, local inflammatory effects, or other mechanisms). The temporary cessation of ovulation has been suggested as a common underlying mechanism behind the reduced risks associated with both the oral contraceptive pill and with parity and breastfeeding - all of which may stop ovulation.^{140,141} This hypothesis could also explain the observation that both younger age at menarche and older age at menopause appear to be associated with increased risks of ovarian cancer, as both these factors would serve to increase a woman's total lifetime number of ovulations.

The potential role of hormones as drivers of ovarian cancer risk is suggested by the observed increased risk amongst women who used menopausal hormones (HT) and amongst women with increased BMI. In particular, the Collaboration found that the association between BMI and ovarian cancer risk was restricted to women who had never used menopausal hormones; one possible explanation is that, in post-menopausal never-users of HT, adipose tissue acts as the main source of oestrogens, and thus a potential driver of ovarian cancer risk, whereas this is likely to be altered amongst users of HT.¹⁷

The apparent reduction in risk of ovarian (and other) cancers associated with long-term use of aspirin and other NSAIDS has been hypothesised to be due to a consequent reduction in chronic inflammation, which may predispose to cancer development.¹⁴² However, an observational analysis in the PLCO trial cohort did not find statistically-significant associations between regular aspirin use and a range of inflammatory markers, after adjustment for multiple comparisons.¹⁴³ Several authors have suggested that local inflammation may have an important role in the development of ovarian cancer, and that this might be a common mechanism behind both the putative reduction in risk with use of anti-inflammatory drugs, and the (still somewhat controversial) possible increase in risk with use of talc, as well as other pro-inflammatory conditions (e.g. endometriosis and pelvic inflammatory disease).^{144–146}

In summary, an understanding of the risk factors for ovarian cancer can be helpful when considering the possible mechanisms by which ovarian cancer may arise and develop. However, it is of increasing importance that epidemiological analyses of the risk factors for ovarian cancer are conducted in prospective studies large enough to perform subgroup analyses by histotype. This explains the rationale for this thesis.

CHAPTER 4

THE MILLION WOMEN STUDY & METHODS

4.1 INTRODUCTION

The main aim of the Million Women Study (MWS) is to investigate the association between reproductive and lifestyle factors and health outcomes amongst (predominantly post-menopausal) women in the UK.

This chapter describes the methods of the Million Women Study, and summarises the approach to statistical analyses used in Chapters 5 to 7.

4.2 DESIGN AND METHODS

The MWS is a multi-centre, population-based, prospective, observational cohort study with continuing follow-up. The methods have been described in detail elsewhere,^{147,148} and are summarised below. Women were asked to participate at the time of invitation to breast screening (mammogram); those who gave consent then completed a questionnaire with information pertaining to a range of reproductive and lifestyle factors, and were followed up using routine cancer registry and mortality data.

Participants are asked to complete a ‘re-survey’ questionnaire every 3-5 years, to update information on exposures (either where the exposure status may have changed over time, e.g. menopausal hormone (HT) use, or where there is now interest from researchers in an exposure that was not asked about on the original questionnaire, e.g. talc use). Copies of all the questionnaires can be viewed on the MWS website at <http://www.millionwomenstudy.org>; a copy of the recruitment questionnaire is in Appendix A.

The study was originally designed primarily to look at the association between use of HT and breast cancer risk, but the wide range of prospectively-collected exposure data, and routinely-collected outcome data, means that the dataset is suitable for exploration of risk factors for many diseases. It is also possible to link the exposure and outcome data to hospital records and, for about 8% of the population, to primary care data.

4.3 RECRUITMENT

4.3.1 Data collection

Women were recruited to the study between 1996 and 2001, making use of the system of routine breast cancer screening within the UK. Between the ages of 50-64, all women in the UK who are registered with a GP (general practitioner) were invited at that time to have routine breast cancer screening by mammogram, at approximately 3-yearly intervals. There are about 100 NHS screening centres in the UK, sixty-six of which agreed to collaborate with the MWS.

During the period of recruitment, participating screening centres posted women the normal screening invitation accompanied by a questionnaire for MWS (in the same envelope). Women were asked to complete the questionnaire and return it at their mammogram appointment. The questionnaire contained questions on a variety of reproductive

and lifestyle factors, together with past medical history, and detailed questions regarding current or previous use of menopausal hormone therapy and the oral contraceptive pill. Participants gave written consent for follow-up, and were asked to give identifying information (to allow linkage with follow-up data). The study includes about one in every four women of eligible age in the UK during the recruitment period.

4.3.2 Data entry and storage

The data from the questionnaires were entered and coded by a combination of optical character recognition software and trained staff, who also performed detailed consistency and validation checks, with reference back to the original questionnaires where necessary.¹⁴⁷ The data are stored, maintained and analysed on a secure site in Oxford, and may not be taken off-site.

4.4 ETHICS

Ethics approval for the Million Women Study was granted by the Oxford and Anglia Multi-Centre Research Ethics Committee (MREC 97/01) (now known as the East of England Cambridge South Research Ethics Committee). Measures have been taken to maintain data confidentiality, including stripping of personal identifiers (which are required to link exposure and outcome data) from the main database used for data analysis. The analyses contained in this thesis are all in keeping with the original aims and restrictions of the original Ethics approval for the MWS.

Access to hospital records was approved by the Information Centre for Health and Social Care (England) and the Information Services Division (Scotland).

4.5 FOLLOW UP

Women who agreed to participate were ‘flagged’ for cancer incidence and cause-specific mortality with the National Health Service Central Register (NHSCR), which is administered by the Health and Social Care Information Centre (HSCIC) (in England), and the National Records of Scotland (in Scotland). [Note that the NHSCR was formerly administered by the Office for National Statistics (ONS), and some references to this persist]. Mortality data come from ONS, via HSCIC.

Hospital admissions data, where relevant, are obtained via the NHS Hospital Episode Statistics in England¹⁴⁹ (HES, available from 1 April 1997 to 31 March 2015), and Scottish Morbidity Records in Scotland¹⁵⁰ (SMR, available from 1 January 1981 to 31 December 2015). HES data were obtained through HSCIC; SMR data were obtained through Information Services Division (ISD) Scotland. Information on diagnoses and procedures associated with each hospital admission were coded to the World Health Organization’s International Classification of Diseases, 10th revision (ICD-10)¹⁵¹ for diagnoses (ICD-9¹⁵² was used up to April 1995) and to the Office of Population Censuses and Surveys classification of surgical operations and procedures, fourth revision (OPCS-4)¹⁵³ for procedures (OPCS-3 was used up to 1988).

Participants contributed follow-up time from the date of recruitment until the earliest of the following outcomes: either diagnosis of the cancer of interest (usually ovarian cancer), diagnosis of another cancer (other than non-melanoma skin cancer), loss to follow-up, death, or the end of follow-up for cancer incidence.

4.6 EXCLUSIONS

The database managers of the Million Women Study have a series of standard exclusions, including women with unsigned forms ($n=114$), and women with missing information on age at recruitment, or who were aged <48 at recruitment ($n=1,775$). Women are also excluded if they cannot be matched with a record at the Office for National Statistics ($n=42$), if they were diagnosed with invasive cancer (other than non-melanoma skin cancer, ICD-10 code C44) prior to recruitment ($n=44,775$), if they have missing information on follow-up date ($n=81$), or if their date of censoring at ONS was prior to recruitment ($n=60$) (see Figure 4.1).

Note that the standard exclusions changed slightly at the most recent database update, which is why numbers in the cohort from older analyses (see Chapters 5 and 6) appear to differ slightly from numbers in more recent analyses (see Chapter 7).

For analyses of ovarian cancer incidence, women were also excluded if they reported a previous bilateral oophorectomy (or were unsure, or the information was missing) ($n=170,772$). Women were also excluded if the exposure status was unknown for the exposure of interest for a given analysis (e.g. for an analysis of tubal ligation and risk of ovarian cancer, women were excluded if their tubal ligation status was unknown).

4.7 OUTCOME OF INCIDENT OVARIAN CANCER

Cases of ovarian cancer were defined as code C56 according to the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10).¹⁵¹ Cases were defined as women who developed ovarian cancer after the date of recruitment and before the latest date of follow-up for their respective study centre.

At the time of analysis, the end date of follow-up for incident cancer for all regions

was 31st December 2014; the end date of follow-up for mortality was 31st December 2015.

For analyses that have already been published at the time of writing this thesis, on tubal ligation and risk of ovarian and other cancers (see Chapters 5 and 6), the follow-up date for incident cancer is 31st December 2013, and the date of follow-up for mortality is 31st December 2014.

4.7.1 Ovarian cancer by histological type

For analyses split by histological subtype, the cases of ovarian cancer were divided into groups based on the coding system of the International Classification of Diseases of Oncology. The cancer registries supplied data based on a mixture of the second¹⁵⁴ and third¹⁵⁵ editions, depending on the date of diagnosis. I used information on histological type from the National Cancer Intelligence Network/ Public Health England where available (NCIN/PHE, available up to December 2013 for England only, and provided as a four-digit code), supplemented by information from the ONS/HSCIC (available for both England and Scotland up to the end of follow-up, and provided as a five-digit code).

For the majority of analyses, cases of ovarian cancer were split into the four most common histotypes: serous (ICD-O codes 8441-8442, 8451, 8460-8463, 9014), mucinous (ICD-O codes 8470-8490), endometrioid (ICD-O codes 8380-8381, 8560, 8570), and clear cell (ICD-O codes 8310, 8313); the remainder of cases were classified as other/ unspecified.

For some analyses, the serous and mucinous tumours were further sub-divided, as previously described.¹⁵⁶ Serous tumours were divided into low-grade tumours (borderline serous tumours (ICD-O codes 8442, 8451, 8462, 8463) or grade 1 serous carcinomas) versus high-grade (grade ≥ 2) serous carcinomas. Information on tumour grade came from NCIN/ PHE data, and thus was not available for all cases. Mucinous tumours

were split into mucinous borderline tumours (ICD-O codes 8472, 8473) and mucinous carcinoma. The coding system used is illustrated in Table 4.1 (note that this table only covers division by ICD-O code, and does not cover division of serous tumours by grade).

4.8 STATISTICAL ANALYSES

Each of the individual results chapters includes a description of the methods used; some common features of the statistical analyses are outlined here.

4.8.1 Descriptive statistics

Simple descriptive statistics (e.g. means and percentages) were used to compare associations between various reproductive and lifestyle factors and both the main exposure (e.g. tubal ligation or parity) and the outcome (e.g. ovarian cancer). For each analysis of a given variable, women were excluded if they had missing information on that variable. Factors which were associated with both the main exposure of interest and the outcome were considered to be possible confounders. Other factors which have been treated as likely confounders by other authors were also considered, even if they did not appear to be associated with either exposure or outcome in this dataset.

4.8.2 Multivariate regression models

For prospective analyses of risk factors for incident cancer, Cox proportional hazards models were used to estimate adjusted hazard ratios (referred to as relative risks (RRs)) of incident cancer according to the exposure of interest. Attained age was the underlying time variable. Graphical methods and tests based on Schoenfeld residuals were used to check for violation of the proportional hazards assumption.

All analyses were stratified by geographical region (10 regions corresponding to the areas covered by the cancer registries: Oxford, East Anglia, South West, Thames, West Midlands, Northern and Yorkshire, Trent, North West (Mersey), North West (Manchester/Lancashire), and Scotland). Analyses of ovarian cancer risk were also generally adjusted for use of the oral contraceptive pill (never, ever), tubal ligation (no, yes), family history of breast cancer (no, yes), hysterectomy (no, yes), use of menopausal hormones (never, ever), body mass index (BMI) (<25 kg/m², 25-29 kg/m², 30+ kg/m²), smoking history (never, past, current), and quintiles of socioeconomic status based on the Townsend deprivation index¹⁵⁷ (calculated using post-code and 1991 census data). For adjustment variables, missing values were assigned to a separate category. All information on adjustment variables was taken from the recruitment questionnaire. Exposure information was either missing or reported as unknown for $<5\%$ of women for all potential confounders.

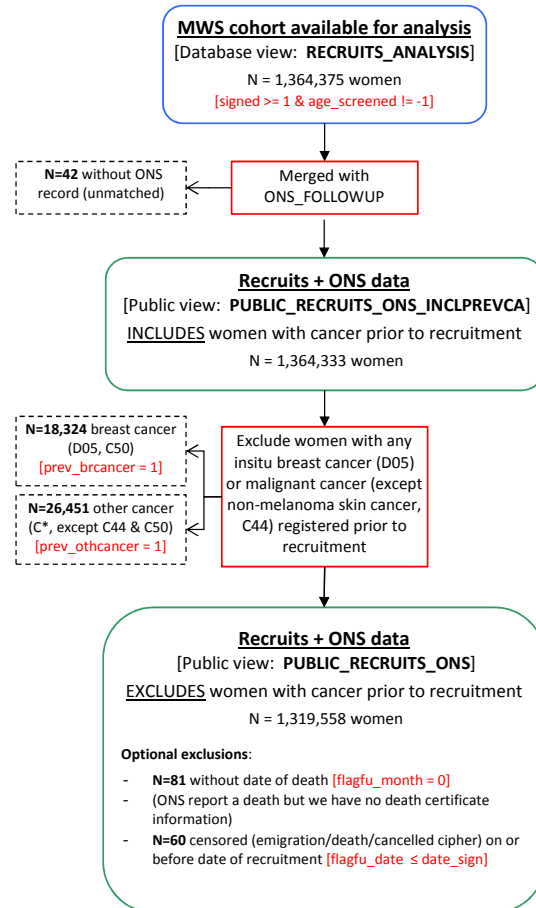
Other factors (including alcohol consumption, physical activity, and age at menarche) were explored as potential confounders, but were generally not included in the final models as their inclusion made no appreciable difference to the main estimate of effect.

Tests of heterogeneity in the relationship between ovarian cancer and the primary exposure by histotype were performed using a competing risks approach.² This method involves the duplication of the data by the number of outcomes (e.g. the number of histotypes), and allows for testing of heterogeneity between multiple outcomes in prospective analyses. This makes it a preferable alternative to multinomial (polytomous) logistic regression, which allows for testing of heterogeneity in the association between an exposure and multiple outcomes, but does not account for variable duration of follow-up. Simulation studies have also suggested that the competing risks approach leads to less biased estimates than multinomial logistic regression when there is violation of the proportionality in hazards functions across disease subtypes.¹⁵⁸

For analyses where comparisons were made between more than two exposure categories, relative risks are presented in figures with group-specific confidence intervals for the log risk in each group (allowing comparisons to be made between any two categories, even if neither is the reference group).^{159,160} Conventional 95% confidence intervals are given in the text. All analyses were performed in Stata-14.¹⁶¹

4.8.3 Plotting of figures

Figures were drawn in R,¹⁶² using either the standard package ‘ggplot,’ or Matthew Arnold’s “Jasper” package.¹⁶³ The size of the markers for the point estimates is proportional to the inverse of the variance (i.e. the square of the standard error). Bars to either side of the markers reflect the 95% confidence intervals. Most figures show relative risk on a log-scale.



Adapted from a figure created by Kirstin Pirie, and used with permission.

Figure 4.1: Flowchart showing detailed standard exclusions for the Million Women Study

Table 4.1: *Classification of the histological types of ovarian cancer*

Histology group	Diagnosis from ICD-O2 or ICD-O3
Serous borderline tumours	
8442 (84421-84425)	Serous cystadenoma, borderline malignancy
8451 (84513-84515)	Papillary cystadenoma, borderline malignancy
8462 (84621-84625)	Papillary serous cystadenoma, borderline malignancy
8463	Serous surface papillary tumor of borderline malignancy
Serous carcinoma	
8441 (84410-84419)	Serous cystadenocarcinoma
8460	Papillary serous cystadenocarcinoma
8461 (84603-84613)	Serous surface papillary carcinoma
9014 (90143)	Serous adenofibroma (malignant version with 3 suffix)
Mucinous borderline tumours	
8472 (84721-84725)	Mucinous cystadenoma, borderline malignancy
8473 (84733, 84735)	Papillary mucinous cystadenoma, borderline malignancy
Mucinous carcinoma	
8470 (84703-84713)	Mucinous cystadenocarcinoma NOS
8471	Papillary mucinous cystadenocarcinoma
8480 (84803-84906)	Mucinous adenocarcinoma
8481	Mucin-producing adenocarcinoma
8490	Signet ring cell carcinoma
Endometrioid	
8380-8381 (83803-83819)	Endometrioid carcinoma / adenofibroma (malignant version with 3 suffix)
8560 (85603)	Adenosquamous carcinoma
85703	Adenocarcinoma with squamous metaplasia

Clear cell	
8310 (83100-83133)	Clear cell adenocarcinoma NOS
8313	Clear cell adenofibroma (malignant version with 3 suffix)
Other	
<i>Epithelial tumours of Other/Mixed/ Not Otherwise Specified (NOS) type</i>	
8010/8230 (80100/82306)	Epithelial tumours/ carcinomas
8240 (82403)	Carcinoid tumour NOS
8243 (82433)	Goblet cell carcinoid
8244 (82443)	Composite carcinoid
8246 (82463)	Neuroendocrine carcinoma, NOS
8260 (82603, 82606)	Papillary adenocarcinoma
8290 (82903)	Oxyphilic adenocarcinoma
8320 (83203)	Granular cell carcinoma
8323 (83233)	Mixed cell adenocarcinoma
8340 (83403)	Papillary carcinoma, follicular variant
84013	Apocrine adenocarcinoma
8440 (84403)	Cystadenocarcinoma NOS
8450 (84503)	Papillary cystadenocarcinoma NOS
85503	Acinar cell carcinoma
8933 (89333)	Adenosarcoma
8940	Mixed tumour, malignant NOS
8950 (89503)	Müllerian mixed tumour
8980 (89803/89809)	Carcinosarcoma NOS
9000 (90003)	Brenner tumour, malignant
9013	Adenofibroma NOS
<i>Non-epithelial tumours</i>	
8620	Granulosa cell tumour, malignant
8631	Sertoli-Leydig cell tumour
8650	Leydig cell tumour, malignant

CHAPTER 4. THE MILLION WOMEN STUDY & METHODS

8800	Sarcoma NOS
8810	Fibrosarcoma NOS
(86203-88103)	As above (grouped)
8854 (88543)	Pleomorphic liposarcoma
8890 (88903)	Leiomyosarcoma NOS
89403	Mixed tumour, malignant NOS
8951 (89513)	Mesodermal mixed tumour
8963 (89633)	Rhabdoid sarcoma
9064	Germinoma
9080	Teratoma, malignant NOS
9084	Teratoma with malignant transformation
9090	Struma ovarii, malignant
90643/90903	As above (grouped)
9110 (91103)	Mesonephroma, malignant
<i>Malignant tumours NOS</i>	
8000-8004 (80000-80043)	Tumour cells, malignant; malignant tumour of small cell/ giant cell/ fusiform cell type

Notes: This table uses a mixture of four-digit and five-digit morphology codes, as the data from ONS came with five-digit codes, while that from NCIN/PHE came with four-digit codes (i.e. omitting the final ‘behaviour’ suffix code).

CHAPTER 5

TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

Please note that much of the material in this chapter has been published in a paper in the International Journal of Cancer¹⁵⁶ (see Appendix C.1).

5.1 INTRODUCTION

Tubal ligation, also known as female sterilization, is a surgical procedure in which the fallopian tubes are occluded (generally by clipping, cutting, or removing a portion of the tube), for the purpose of permanent contraception.

5.1.1 Tubal ligation: trends with age and time

Data from the annual reports on Contraception and Sexual Health (by the Office for National Statistics) show that female sterilization has become less common over time in the UK: while 11% of women aged 16-49 reported having undergone female sterilization in the survey of 2000-2001, this had dropped to 6% in the survey of 2008-2009.¹⁶⁴ Although

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

tubal ligation has become less common as a form of contraception over time, there has remained a strong trend with age, with tubal ligation being very rare at younger ages, and more common at older ages (see Figure 5.1). This trend presumably relates to the fact that tubal ligation is usually irreversible, and so should only be performed after a woman is certain that she does not want to have any more children - and hence tends to be performed later in reproductive life.

These trends of tubal ligation prevalence with age and time mean that, if one wishes to explore epidemiological associations with tubal ligation, it is easier to do so in populations of older women, who had their reproductive years longer ago (when tubal ligation was more common).

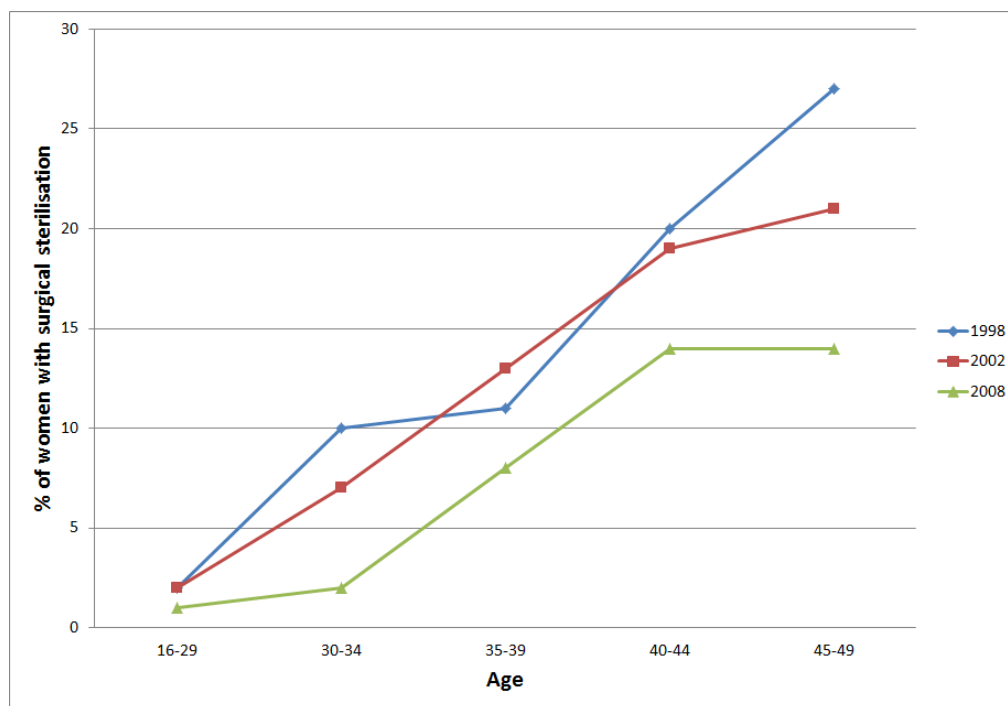


Figure 5.1: Graph showing percentage of women reporting female sterilization (tubal ligation) by age and year (Source: Office for National Statistics^{164–166})

5.1.2 Tubal ligation: associations with ovarian cancer

Several meta-analyses of observational epidemiological studies have found that tubal ligation is associated with an overall reduced risk of ovarian cancer.^{1,127,167} Pooled individual-patient analyses of 13 prospective²⁷ and 13 retrospective¹ studies both found about a 20% reduction in risk of ovarian cancer amongst women with tubal ligation; meta-analyses of published studies estimated about a 30-35% reduction in risk.^{127,167} Published findings by tumour histotype are inconsistent - except perhaps that most investigators have reported a reduced risk for endometrioid tumours.^{1,27,127,167}

However, the majority of studies to examine this association, and to have enough cases to explore possible heterogeneity by histological type, have been retrospective studies. As discussed previously (see Chapter 3), there is reason to believe that the greater potential for bias in retrospective studies may sometimes lead to different (and probably incorrect) findings, compared to prospective studies.

Epidemiological studies need to be sufficiently large to have enough power to reliably detect heterogeneity by tumour histotype. This chapter therefore reports on the association observed between tubal ligation and incident ovarian cancer in a large cohort study of UK women.

5.2 AIM

The aim of this study is to examine the association between tubal ligation and risk of incident ovarian cancer in a prospective study, and to explore possible heterogeneity of this association by histological tumour type.

5.3 METHODS

The methods of recruitment and follow-up for the Million Women Study have been described in detail in Chapter 4, and are summarised below.

5.3.1 Study design, data collection and follow-up

The Million Women Study is a prospective study of 1.3 million UK women, recruited in 1996-2001 via the UK National Health Service (NHS) Breast Screening Programme. At recruitment, women completed a questionnaire on socio-demographic, reproductive, medical and lifestyle factors. The cohort is resurveyed approximately every three to five years. The study design and methods are described in detail elsewhere,^{147,148} and questionnaires can be viewed online at <http://www.millionwomenstudy.org> (also see Appendix A).

All participants have been flagged on the NHS Central Register (NHSCR), so the study investigators are routinely notified of cancer registrations and deaths. The information provided includes the date of the event (cancer registration or death), together with the cancer site (coded using the 10th revision of the International Classification of Diseases, ICD-10)¹⁵¹ and tumour morphology (coded using the second and third editions of the International Classification of Diseases for Oncology, ICD-O).^{154,155} The Million Women Study has also been linked to data from the National Cancer Intelligence Network (NCIN), providing additional information on diagnostic histology codes, and on tumour grade, for cases diagnosed in England up to 31st December 2013. For this study, tumour histology and grade information from NCIN were used where available, to supplement the NHSCR data.

All participants gave written consent to follow-up at recruitment. Ethical approval was granted by the Oxford and Anglia Multi-Centre Research Ethics Committee (MREC

97/01).

5.3.1.1 Exposure variables

Women in the study were asked at recruitment ‘Have you been sterilized (had your tubes tied)?’, and if so, their age at sterilization (which was used for analyses of the timing of sterilization).

5.3.1.2 Outcome

The outcome of interest was ovarian cancer (C56 in ICD-10).¹⁵¹ For histotype analyses, the outcome was split into five histological groups: serous (ICD-O codes 8441-8442, 8451, 8460-8463, 9014), mucinous (ICD-O codes 8470-8490), endometrioid (ICD-O codes 8380-8381, 8560, 8570), clear cell (ICD-O codes 8310, 8313), and other.

For some exploratory analyses, tumour types were analysed separately by grade (available for about 40% of all cases). Three-tier grading systems had generally been used (including those outlined by FIGO (International Federation of Gynecology and Obstetrics),¹⁶⁸ WHO (World Health Organization),¹⁶⁹ and Silverberg¹⁷⁰), and the few cases classified as ‘grade 4’ were grouped together with ‘grade 3’ tumours.

Serous tumours were divided into low-grade tumours (defined here as serous borderline tumours (ICD-O codes 8442, 8451, 8462, 8463) or grade 1 serous carcinomas), and high-grade serous carcinomas (defined here as grade ≥ 2 serous carcinomas). The assumptions made in converting tumour grade information from the three-tier (grade 1-3) to the two-tier (low-grade vs high-grade) system for serous carcinoma were in line with those used in other publications.^{23,54,171}

Endometrioid tumours were not divided by grade, as there was not a strong *a priori* reason to think that risk factors would be different for low-grade vs high-grade

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

tumours (personal communication with Dr Sanjiv Manek, consultant histopathologist), and numbers were relatively small.

Clear cell tumours were not divided by grade, as all clear cell ovarian cancers are high-grade / grade 3, by definition.^{23,172,173}

Mucinous tumours were split into mucinous borderline tumours (ICD-O codes 8472-8473) and mucinous carcinoma (ICD-O codes 8470-8471 and 8480-8490); the mucinous carcinomas were not further subdivided by tumour grade, as numbers were small.

5.3.2 Statistical analyses

As part of standard exclusions, women were excluded if they had missing information on date of birth (n=29) or follow-up (n=61), or if the date of censoring at ONS was before recruitment (n=57). Women were also excluded from these analyses if: (1) they had been diagnosed with any invasive cancer other than non-melanoma skin cancer (ICD10 code C44) prior to recruitment (n=44,829), (2) they reported at recruitment having had both ovaries removed (bilateral oophorectomy), or if they were unsure whether they had or not (n=170,769), or (3) they had missing data on tubal ligation (n=15,712) (see Figure 5.2).

The remaining women (N=1,132,914) contributed person-years from the date of recruitment in to the study until the date of registration for ovarian cancer, the date of death, or last date of follow-up (31st December 2013) – whichever was soonest. Women were censored at diagnosis of any non-ovarian cancer (other than non-melanoma skin cancer). For analyses exploring effects of the timing of tubal ligation, women were excluded if they had missing information on age at tubal ligation (n=14,684). About 1% of participants had been lost to follow-up and such women are censored at the date when they were lost, contributing person-years until then.

Cox (proportional hazards) regression models were used to estimate hazard ratios (referred to as relative risks (RRs)) of developing ovarian cancer by tubal ligation status. Attained age was the underlying time variable. There was no evidence of significant violation of the proportional hazards assumption, as assessed by graphical methods and tests based on Schoenfeld residuals.

All analyses were stratified by geographical region (10 regions corresponding to the areas covered by the cancer registries), and further adjusted for parity (0, 1, 2, 3+), use of the oral contraceptive pill (OCP) (never, ever), family history of breast cancer (no, yes), hysterectomy (no, yes), use of menopausal hormones (never, ever), body mass index (BMI) (<25 kg/m², 25-29 kg/m², 30+ kg/m²), smoking history (never, past, current), and quintiles of socioeconomic status (based on the Townsend deprivation index).¹⁵⁷ All adjustment variables were as reported at recruitment. For adjustment variables, missing values were assigned to a separate category. Exposure information was either missing or reported as unknown for $\leq 6\%$ of women for all potential confounders. Sensitivity analyses were conducted excluding all women with missing data on covariates within the model.

Other factors (including alcohol consumption, physical activity, and age at menarche) were explored as potential confounders, but were not included in the final model as their inclusion made no appreciable difference to the main estimate of effect.

Tests of heterogeneity in the relationship between tubal ligation and ovarian cancer risk by histotype were performed using a competing risks approach.²

To explore possible effect modification of the main relationship between tubal ligation and ovarian cancer, multiplicative interaction terms were modelled between tubal ligation and each potential confounder in turn. Likelihood ratio tests were used to compare the models with versus without the interaction terms. For these comparisons, analyses were limited to participants without missing information on the potential effect-modifier

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

of interest, and so numbers of cases will differ slightly.

Analyses were also conducted to explore possible effects of age and time with tubal ligation, including age at, time since, and calendar year of tubal ligation, and whether the tubal ligation was in the same year as the last birth or afterwards. All analyses of the timing of tubal ligation were restricted to parous women, as there were very few nulliparous women who had had a tubal ligation. For these analyses (where comparisons are made between more than two exposure categories), relative risks were treated as floated absolute risks and are presented in figures with group-specific confidence intervals for the log risk in each group (allowing comparisons to be made between any two categories, even if neither is the reference group).¹⁵⁹ Conventional 95% confidence intervals (CI) are given in the text.

Analyses were performed in Stata-14.¹⁶¹ Tests of statistical significance were two-sided. Figures were drawn in R¹⁶² using Matthew Arnold's "Jasper" package.¹⁶³

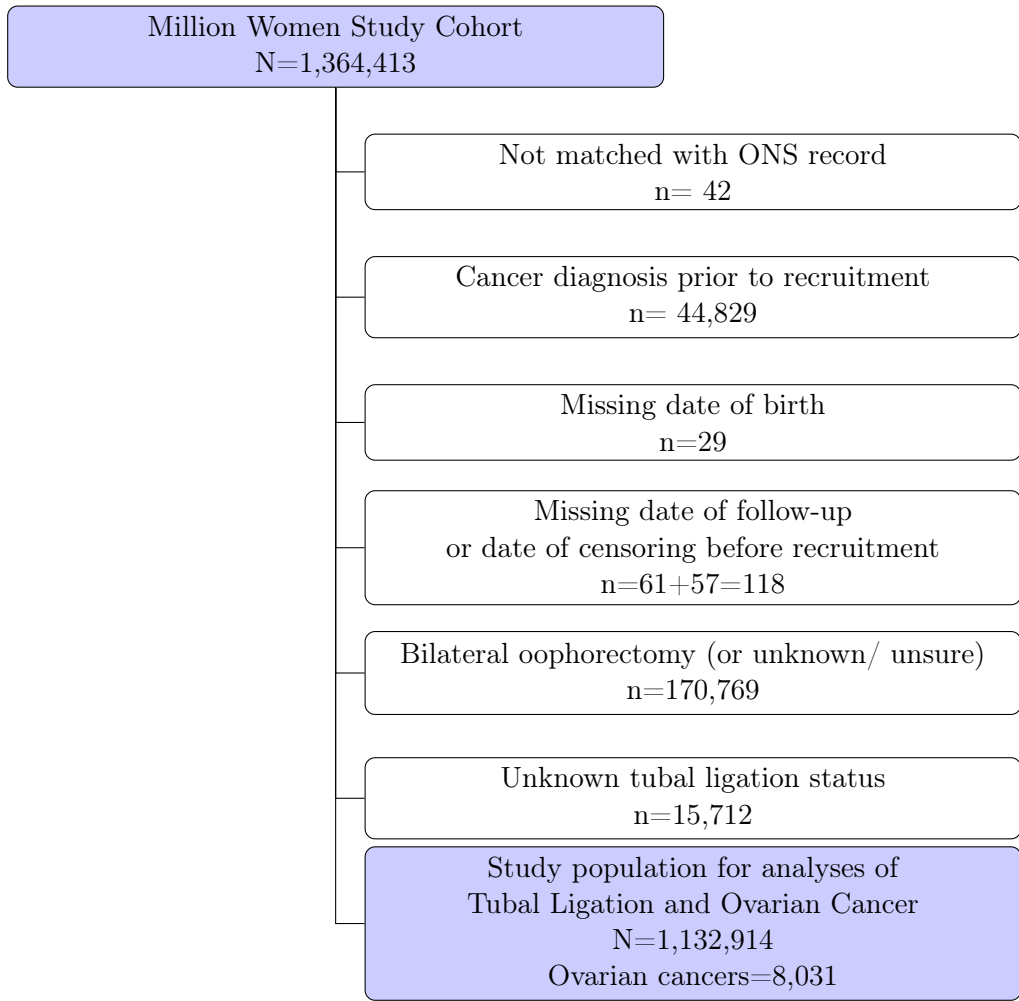


Figure 5.2: *Participants inclusion/exclusion flow diagram for analyses of tubal ligation and ovarian cancer*

5.4 RESULTS

A total of 1,132,914 women, mean age at recruitment 56.1 (standard deviation (SD) 4.8) years, were included in the analyses. At recruitment, 246,048 (22%) reported having had a tubal ligation, at median age 35 (interquartile range (IQR) 31-38), and median year 1978 (IQR 1973-1982). Women who reported a previous tubal ligation were less likely to be nulliparous (3% vs 13%), and were more likely to have used the oral contraceptive

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

pill (70% vs 57%) and menopausal hormones (55% vs 45%), to have had a hysterectomy (19% vs 14%), to be a current smoker (26% vs 18%), and to live in areas in the lower third of socioeconomic status (38% vs 31%), compared to women without tubal ligation (see Table 5.1).

The women were followed up for incident ovarian cancer over 15.6 million person-years, with a mean duration of follow-up of 13.8 (SD 3.4) years per woman. During this period, 8,031 incident ovarian cancers were registered, of which 3,517 (44%) were serous; 836 (10%) were mucinous; 590 (7%) were endometrioid; 401 (5%) were clear cell; and 2,687 (33%) were of other histological types (mostly unspecified epithelial tumours). The mean age at diagnosis of ovarian cancer was 65.1 years (SD 6.4).

The relative risk of ovarian cancer amongst women with tubal ligation compared to those without was 0.80 (95% CI: 0.76-0.85, $p < 0.001$), after adjustment for age, region, parity, family history of breast cancer, hysterectomy, socioeconomic status, body-mass index, smoking, and use of contraceptive or menopausal hormones.

5.4.1 Variation by histotype

There was strong evidence of heterogeneity by histotype (see Figure 5.3; heterogeneity, $p = 0.000002$). Tubal ligation was associated with about a halving in the risk of endometrioid (RR: 0.45, 95% CI: 0.34-0.59) and clear cell tumours (RR: 0.55, 95% CI: 0.39-0.77), and a lesser but still significant reduction in the risk of serous tumours (RR: 0.84, 95% CI: 0.77-0.92), but there was no significant reduction in the risk of mucinous tumours (RR: 0.99, 95% CI: 0.84-1.18).

When serous tumours were divided into low-grade serous tumours (serous borderline tumours and low-grade serous carcinoma) and high-grade serous carcinoma, risks differed significantly (heterogeneity, $p = 0.01$). There was strong evidence that tubal ligation was

associated with a reduction in the risk of high-grade serous carcinoma (n=1,682, RR: 0.80, 95% CI: 0.70-0.91), but not of low-grade serous tumours (n=447, RR: 1.12, 95% CI: 0.89-1.42).

When mucinous tumours were split into borderline and fully-malignant subtypes, there was little or no association between tubal ligation and the risk of either mucinous borderline tumours (n=411, RR: 0.98, 95% CI: 0.76-1.25) or fully-malignant mucinous carcinoma (n=425, RR: 1.01, 95% CI: 0.79-1.30).

A sensitivity analysis excluding all women with missing values in any of the adjustment variables showed that the association between tubal ligation and ovarian cancer, and the variation by histology, was not appreciably changed (data not shown).

5.4.2 Timing of tubal ligation

There was little or no variation in risk either by age at, year of, or years since tubal ligation, or whether tubal ligation was performed in the same year as the last birth or subsequently, for ovarian cancer overall, and for endometrioid and clear cell tumours (see Figure 5.4).

Compared to women without tubal ligation, for serous tumours there was some suggestion of a greater reduction in risk with tubal ligations performed ≤ 1974 (RR: 0.69, 95% CI: 0.59-0.81), than with those performed in 1975-1979 (RR: 0.93, 95% CI: 0.80-1.07), or those performed in 1980 or later (RR: 0.84, 95% CI: 0.73-0.97), heterogeneity: $p=0.02$. There was also some suggestion of a greater reduction in risk for women with a time since tubal ligation of >25 years (RR: 0.77, 95%CI: 0.70-0.86) than for ≤ 25 years since tubal ligation (RR: 0.94, 95%CI: 0.80-1.09), heterogeneity: $p=0.04$.

For mucinous tumours, there was a suggestion that the association with tubal ligation varied by age at tubal ligation, and by timing of tubal ligation in relation to the last

birth (see Figure 5.4). However, given that the overall association between tubal ligation and risk of mucinous tumours was null, these apparent heterogeneous associations in subgroup analyses may be due to chance and multiple testing.

5.4.3 Exploring possible effect modification

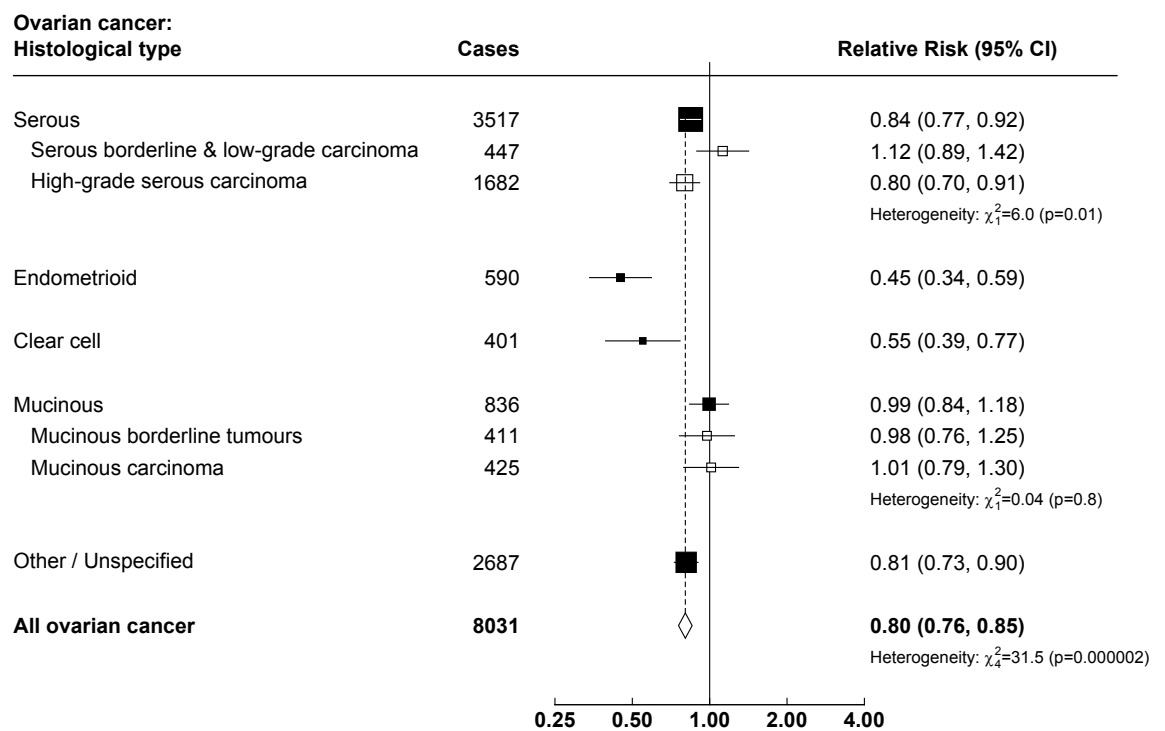
There was no significant variation in the association between tubal ligation and ovarian cancer with parity, use of contraceptive or menopausal hormones, hysterectomy, age at natural menopause, family history of breast cancer, smoking, socioeconomic status, frequency of strenuous exercise, or alcohol intake (see Figure 5.5). There was some weak evidence of heterogeneity by body mass index and age at menarche, although this may well be due to chance, as multiple statistical tests were performed.

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

Table 5.1: *Characteristics of the study population at recruitment, and details of follow-up for ovarian cancer, according to tubal ligation status*

	Tubal ligation		All women
	No	Yes	
Characteristics			
Number of women	886,866	246,048	1,132,914
Mean (SD) age at recruitment (years)	56.3 (4.9)	55.3 (4.3)	56.1 (4.8)
Socioeconomic status, lower third, % (n)	30.5 (268,314)	38.0 (92,663)	32.1 (360,977)
Mean (SD) age at menarche	13.0 (1.6)	13.0 (1.6)	13.0 (1.6)
Nulliparous, % (n)	13.3 (117,319)	2.6 (6,352)	10.9 (123,671)
Ever use of oral contraceptive pill, % (n)	57.0 (501,857)	70.3 (171,546)	59.9 (673,403)
Current use of menopausal hormones, % (n)	44.5 (390,518)	54.9 (133,529)	46.7 (524,047)
Hysterectomy, % (n)	14.2 (126,054)	19.4 (47,579)	15.4 (173,633)
Mean (SD) age at natural menopause	49.3 (4.2)	48.7 (4.5)	49.1 (4.3)
Family history of breast cancer, % (n)	9.8 (82,101)	10.0 (22,781)	9.9 (104,882)
Mean (SD) body mass index (kg/m ²)	26.0 (4.6)	26.6 (4.8)	26.1 (4.7)
Current smoker, % (n)	18.4 (153,773)	26.4 (61,264)	20.1 (215,037)
Strenuous exercise ≥once/week, % (n)	39.8 (340,538)	37.7 (89,266)	39.3 (429,804)
Alcohol intake, ≥7 units/week, % (n)	23.6 (207,870)	25.5 (62,130)	24.0 (270,000)
Follow-up for ovarian cancer incidence			
Woman-years of follow-up for incidence (100,000s)	122.6	33.9	156.5
Mean follow-up time per woman (SD)	13.8 (3.4)	13.8 (3.3)	13.8 (3.4)
Number of incident ovarian cancers	6,689	1,342	8,031
Notes: Means and percentages are calculated excluding missing values for the variable of interest.			
n: number of women. SD: standard deviation.			

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES



Adjusted for age, region, deprivation, parity, hysterectomy, family history of breast cancer, BMI, smoking, and use of contraceptive and menopausal hormones.

Figure 5.3: Relative risk of ovarian cancer histotypes in women with versus without a history of tubal ligation

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

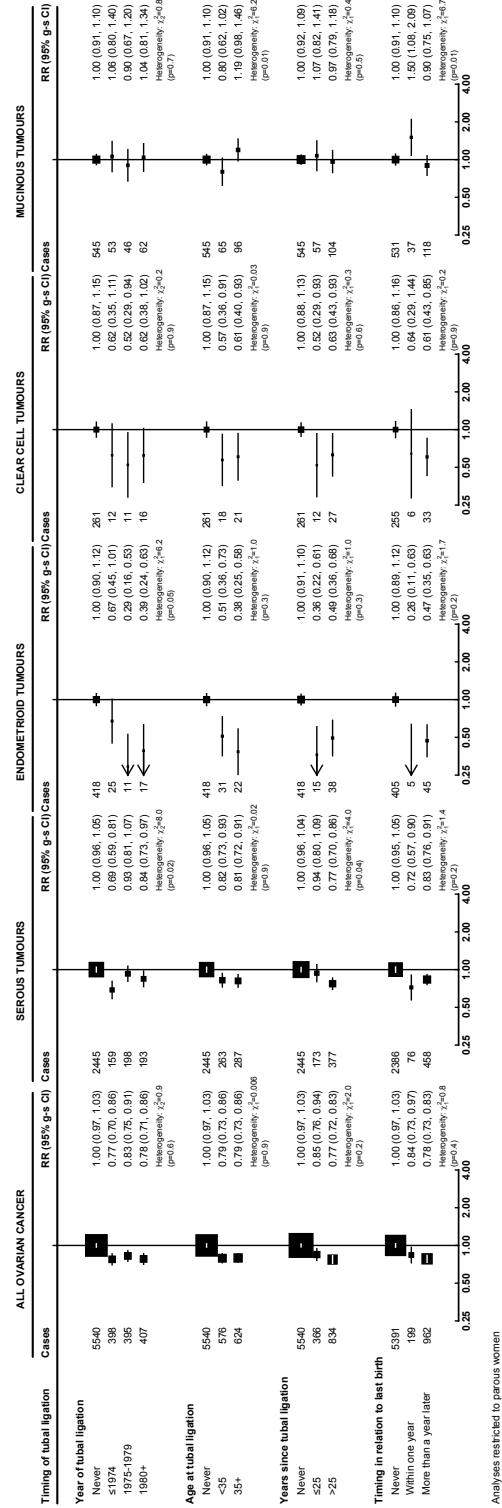


Figure 5.4: Relative risk of ovarian cancer (overall and for the four main histotypes) in relation to the timing of tubal ligation

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

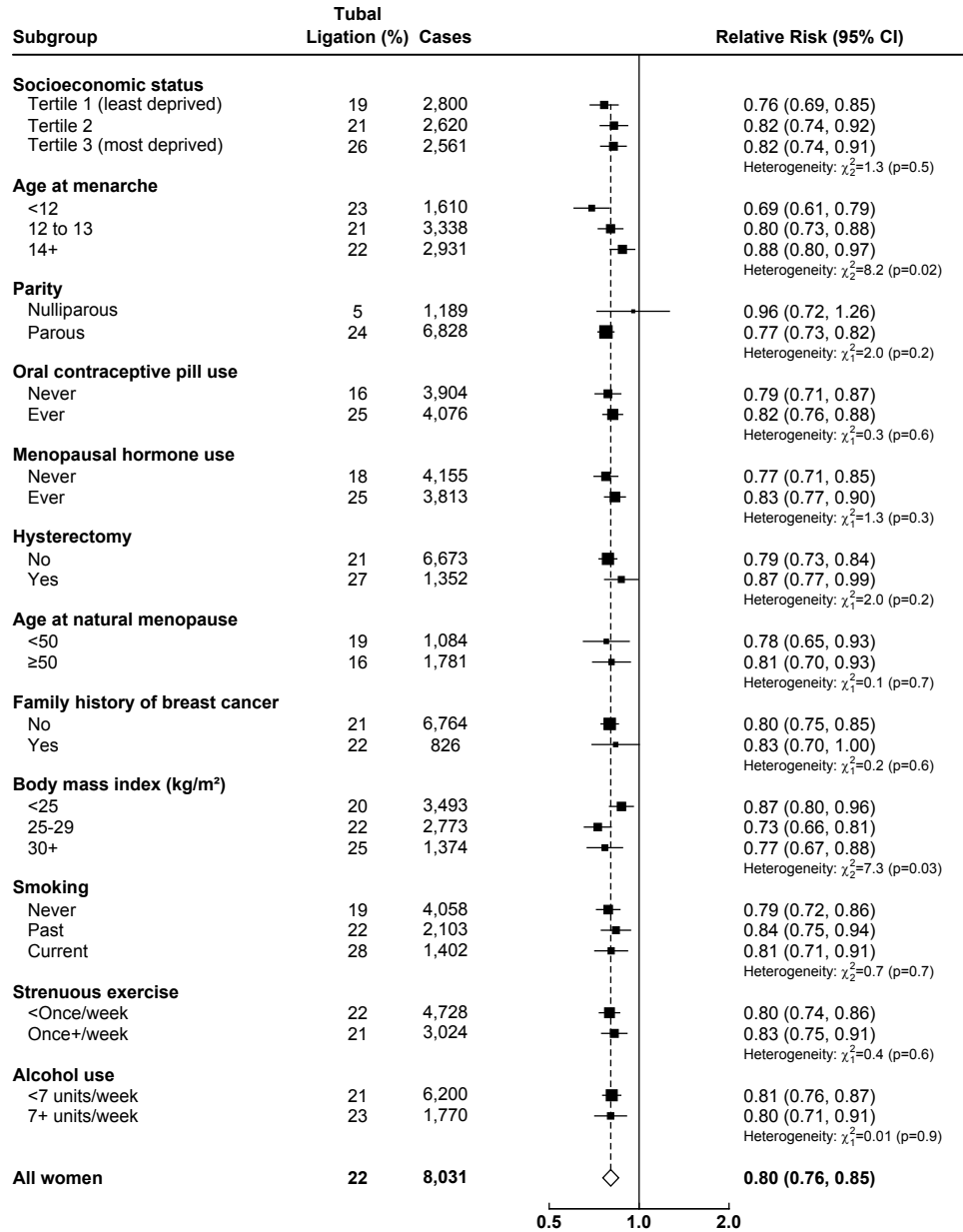


Figure 5.5: Relative risk of ovarian cancer amongst women with versus without tubal ligation, by subgroups

5.5 DISCUSSION

These analyses in a large prospective cohort study with 8,031 incident cases of ovarian cancer showed strong evidence that the risk of ovarian cancer associated with tubal ligation varied by tumour histotype. A modest reduction in risk was found for the most common histotype, serous tumours, as had been previously reported.^{1,127,167} However, when serous tumours were split into low-grade serous tumours and high-grade serous carcinoma, there were significant differences, with a 20% reduction in risk of high-grade serous carcinoma, but no reduction in risk of low-grade serous tumours. This is the first study to report a significant difference between the risks of low-grade versus high-grade serous ovarian tumours associated with tubal ligation.

A pooled analysis of individual-patient data from 13 retrospective studies had reported that tubal ligation was associated with a reduced risk of high-grade serous ovarian cancer (OR: 0.80, 95% CI: 0.73-0.89; n=3,791), but this did not differ significantly from the association with low-grade serous ovarian cancer (OR: 0.89, 95% CI: 0.65-1.22; n=361).¹ A pooled analysis of data from 21 prospective studies (13 of which had information on tubal ligation, with about 4,000 cases of ovarian cancer) did not find a significant reduction in the risk of serous carcinomas overall (RR: 0.91, 95% CI: 0.79-1.06), nor when split into low-grade and high-grade tumours.²⁷ [However, these results are not directly comparable, as the authors of these studies did not choose to group serous borderline tumours with low-grade serous carcinomas.]

Tubal ligation was associated with about a halving of the risk of endometrioid and of clear cell tumours. Others have reported similar findings for endometrioid tumours: although both endometrioid and clear cell tumours are sufficiently rare that relatively few individual studies have reported significant associations with tubal ligation, meta-analyses have been fairly consistent in finding that tubal ligation is associated with about

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

a halving of risk of endometrioid tumours.^{1,27,127,167} Reported findings for clear cell tumours from previous studies are mixed, perhaps because it is relatively uncommon. One meta-analysis did not report separately for clear cell tumours, presumably because numbers were small.¹²⁷ Those meta-analyses that have reported on the association between tubal ligation and clear cell ovarian cancer have tended to find a reduction in risk,^{1,27,167} though sometimes with quite wide confidence intervals.

This analysis found little or no association between tubal ligation and the risk of mucinous ovarian cancer, in keeping with some reports,^{27,127,167} though not others.¹

5.5.1 Mechanisms

New hypotheses about the origins of ovarian cancer have provoked a reconsideration of possible mechanisms underlying the reduced risk of ovarian cancer with tubal ligation.^{1,127,174,175} The significant heterogeneity seen here between low-grade versus high-grade serous tumours is consistent with the hypothesis that the two tumour subtypes develop along distinct pathways, with the high-grade serous carcinomas thought to arise from precursor lesions within the epithelium of the fimbrial end of the fallopian tubes, while the serous borderline tumours and low-grade serous carcinomas may arise as part of a different pathway, possibly involving epithelial inclusion cysts within the ovary²¹ (discussed in more detail in Chapter 2).

Surgical ligation of the fallopian tube might reduce the risk of high-grade serous ovarian cancer, either by physically obstructing the passage of tubal cells towards the ovary, or by collateral damage to the local tubal blood supply, with potential subsequent effects on the tubal epithelium. In addition, some forms of tubal sterilization involve removal of some or all of the tube, and/or disruption of the distal end of the tube – which might have been more common in the past, and could possibly explain the significantly greater risk reduction found for serous tumours for tubal ligations done before 1974.

The apparent lack of an effect of tubal ligation for low-grade serous tumours is consistent with the hypothesis that they have a different origin to high-grade serous ovarian cancer. However, it is not entirely incompatible with a possible tubal origin for low-grade serous tumours (for example, if the precursor cells of origin had already been attached to the ovary before the tubal ligation occurred).

Some endometrioid and clear cell tumours are thought to develop from endometriosis^{22,23} (also discussed in more detail in Chapter 2) and one might speculate that the almost halving of risk seen with tubal ligation could be due to the blocking of retrograde menstruation through the fallopian tubes to the ovaries and pelvis. It could also reflect the blocking of the transport of carcinogens up the tubes to the ovaries and pelvis.

For any mechanism that involved the blockage of the passage of carcinogens or cells (whether from the endometrium or fallopian tube) along the tube, one might expect the association with tubal ligation to vary by timing (e.g. a greater reduction in risk for tubal ligation performed at a younger age, or longer ago). However, we found no suggestion of any association with the timing of tubal ligation for ovarian cancer overall or for endometrioid and clear cell tumours, and only limited evidence of an association with the year of, and years since, tubal ligation for serous tumours. These findings are in keeping with most, but not all, previous studies.^{1,127,167}

5.5.2 Strengths and weaknesses

This study has several strengths, in addition to the large sample size and number of cases. In particular, the prospective collection of exposure data helped to prevent differential recall of tubal ligation or other factors amongst women with and without cancer. The use of routinely-collected national data for follow-up and ascertainment of incident ovarian cancer resulted in few participants being lost to follow-up. The prospective study design also facilitated the involvement of women with more aggressive tumour types, who might

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

not be included in retrospective studies, as there can be a substantial time lag between cancer diagnosis and recruitment to a retrospective study. The high prevalence of tubal ligation (22%) was also an advantage in terms of statistical power. By comparison, a nationwide prospective study in Denmark, with 13,241 cases of ovarian cancer, was relatively under-powered due to the rarity of tubal ligation in that population (around 3%).¹⁷⁵

One weakness of the study is that, despite being based on probably the largest prospective cohort of its kind, it nevertheless became underpowered for some subgroup analyses. This was particularly evident in analyses of the timing of tubal ligation by histological subtype, when there were very few cases of endometrioid or clear cell ovarian cancer in some of the timing subgroups.

Another potential weakness of the study was the absence of central histopathological review of ovarian cancer cases, for standardization and incorporation of the latest diagnostic criteria. The distribution of histotypes amongst the cases in this study was consistent with those from other population-based studies, with serous tumours accounting for around 65% of those of the four main histotypes, endometrioid tumours for 11%, mucinous 15%, and clear cell 7% (by way of comparison, the corresponding percentages from the EPIC cohort are: 67% serous, 13% endometrioid, 14% mucinous, and 6% clear cell¹⁷⁶). This distribution is slightly different to that reported by two hospital-based studies of ovarian cancer cases, in both of which mucinous tumours accounted for only around 3% of cases;^{177,178} this may reflect differences in the study populations (North America vs Europe), or historical differences in diagnostic criteria (for example, mucinous tumours in this cohort might include some metastases from the gastrointestinal tract).

One particular issue with the absence of central histopathological review was regarding the classification of serous carcinomas: although these would traditionally have been graded on a three-tier system (low, medium and high grade), new classification

systems are binary, with low-grade and high-grade serous carcinomas now considered to be different entities. In the absence of central review, we had to assume that Grade 1 serous carcinomas would correspond to low-grade, while Grade 2+ would correspond to high-grade; although these assumptions are in line with other studies, it is nevertheless likely that Grade 2 serous carcinomas include some misclassified low-grade tumours. If so, one might expect this misclassification to mask any true differences in association between low-grade and high-grade serous carcinomas - although in fact heterogeneity was observed between them in their associations with tubal ligation.

33% of ovarian cancers in this study were not recorded as being of one of the four main histotypes (serous, endometrioid, clear cell or mucinous). Of these, 87% were epithelial tumours of other, mixed, or unspecified type (mostly unspecified adenocarcinoma, ICD-O code 8140, or unspecified carcinoma, ICD-O code 8010), 9% were unspecified malignant tumours (ICD-O code 8000), and 4% were non-epithelial tumours. Given that high-grade serous carcinoma is the most common histotype of ovarian cancer, it seems likely that many of the tumours listed as unspecified carcinoma/ adenocarcinoma are in fact high-grade serous tumours. This would be supported by the observation that tubal ligation is associated with a similar reduction in risk of other/unspecified ovarian cancer (RR: 0.81, 95% CI: 0.73-0.90) as of high-grade serous carcinoma (RR: 0.80, 95% CI: 0.70-0.91).

5.5.3 Conclusion

These analyses have confirmed in the largest prospective study to date that tubal ligation is associated with a reduced risk of ovarian cancer overall, with significant heterogeneity by histological tumour type. Tumour histology may have been classified in somewhat different ways by different pathologists. Any such misclassification would tend to blur differences by tumour histotype, yet distinctly heterogeneous risks were found. This

CHAPTER 5. TUBAL LIGATION AND RISK OF OVARIAN CANCER HISTOTYPES

argues strongly for causality, as it confirms that the variation in risk associated with tubal ligation is not due just to confounding, and strengthens hypotheses that different histotypes of ovarian cancer have different causes. In particular, the significant reductions in risk seen with tubal ligation for high-grade serous carcinomas, and the even more substantial reductions in risk for endometrioid and clear cell carcinomas, are consistent with hypotheses that many cases of these histotypes may arise from sites other than the ovary.

CHAPTER 6

TUBAL LIGATION AND RISK OF 26 SITE-SPECIFIC CANCERS

Please note that much of the material in this chapter has been published in a Brief Communication in the British Journal of Cancer,¹⁷⁹ see Appendix C.2.

6.1 INTRODUCTION

As shown in Chapter 5, women who have undergone tubal ligation (female sterilization) are at a reduced risk of ovarian cancer, with an association that varies by histological tumour type. Other investigators have reported on associations between tubal ligation and risk of cancers at other sites, with varied results. For example, associations have been described with breast,^{180,181} endometrial,¹⁸² cervical¹⁸³ and anal cancers.⁵

However, few studies have been able to explore associations with cancers of the peritoneum and fallopian tube (which might be particularly interesting to investigate, as the most common forms of these cancers are high-grade serous tumours, histologically and clinically similar to those of the ovary), as they are diagnosed relatively rarely.

CHAPTER 6. TUBAL LIGATION AND RISK OF 26 SITE-SPECIFIC CANCERS

A few previous reports regarding tubal ligation and cancer risk come from prospective studies, investigating associations between tubal ligation and several different cancer types (for example, a prospective study in Shanghai, which reported on associations between tubal ligation and 11 cancer types,¹⁸⁴ and a register-based study in Denmark, which reported on tubal ligation and the risk of several gynaecological cancers¹⁸²). However, the majority of results come from cancer-specific retrospective studies.

6.2 AIM

The aim of this study is to systematically examine associations between tubal ligation and risk of common or relevant cancers (i.e. any cancer for which we had 500 or more cases, plus fallopian tube cancers) in the Million Women Study. This would be useful not only in obtaining reliable evidence of the true association between tubal ligation and risk of these specific cancers, but would also provide supportive information when considering the possible mechanisms behind the known reduction in risk of ovarian cancer.

6.3 METHODS

The methods of recruitment and follow-up for the Million Women Study have been described in detail in Chapter 4, and are summarised below.

6.3.1 Study design, data collection and follow-up

The Million Women Study is a prospective study of 1.3 million UK women, recruited in 1996-2001, at the time of invitation for breast cancer screening. Participants completed a questionnaire at recruitment, on socio-demographic, reproductive and lifestyle factors. The study design and methods have been described previously^{147,148} and questionnaires can be viewed online at <http://www.millionwomenstudy.org> (also see Appendix A).

Follow-up information for outcomes (cancer registrations and deaths) came from the NHS Central Register (on which all MWS participants are ‘flagged’). The information provided includes the date of registration and cancer site (coded using the 10th revision of the International Classification of Diseases, ICD-10¹⁵¹). All participants gave written consent to follow-up at recruitment. Ethical approval was granted by the Oxford and Anglia Multi-Centre Research Ethics Committee (MREC 97/01).

6.3.1.1 Exposure variable

Information on tubal ligation status was taken from responses to the question ‘Have you been sterilized (had your tubes tied)?’ on the recruitment questionnaire.

6.3.1.2 Outcomes

Analyses were run separately for each of twenty-six cancer types or sites, defined by ICD-10 code, and for all cancer (any invasive cancer other than non-melanoma skin cancer, ICD-10 C44). Analyses were limited to cancer types for which 500 or more cases with an identified primary site had accrued. In addition, an analysis was conducted for fallopian tube cancers (ICD-10 C57.0), despite having fewer than 500 cases, as there is a reasonable prior hypothesis that surgery to the fallopian tube might affect the incidence of cancer at this site.

6.3.2 Statistical analysis

Women were excluded as standard if they did not match with the ONS records (n=42), if they had missing information on date of birth (n=29) or follow-up (n=61), if their date of censoring at ONS was prior to recruitment (n=57), or if they had invasive cancer other than non-melanoma skin cancer (ICD10 code C44) prior to recruitment (n=44,829).

CHAPTER 6. TUBAL LIGATION AND RISK OF 26 SITE-SPECIFIC CANCERS

Women were also excluded from these analyses if they had missing data on tubal ligation ($n=40,612$) (see Figure 6.1). The remaining women ($N=1,278,783$) contributed person-years from the date of recruitment until the earliest of the date of registration of any cancer (other than non-melanoma skin cancer), the date of death, or last date of follow-up (31st December 2013). About 1% of participants were lost to follow-up and contributed person-years until the date they were lost.

Cox (proportional hazards) regression models were used to estimate hazard ratios (referred to as relative risks (RRs)) of developing a given cancer by tubal ligation status. Attained age was the underlying time variable. All analyses were stratified by geographical region (10 regions corresponding to the areas covered by the cancer registries) and quintiles of socioeconomic status (based on the Townsend deprivation index),¹⁵⁷ and additionally adjusted for parity (0, 1, 2, 3+), age at first birth (<20, 20-24, 25-29, 30+), smoking (never, past, current <10 cigarettes/day, current 10-19 cigarettes/day, current 20+ cigarettes/day), average alcohol intake per week (none, ≤ 7 units, >7 & ≤ 14 units, >14 units), frequency of strenuous exercise (<once/week, $\geq$ once/week), use of menopausal hormones (never, ever), use of the oral contraceptive pill (never, ever), hysterectomy (no, yes), and body mass index (<25 kg/m², 25-29 kg/m², 30+ kg/m²).

All adjustment variables were as reported at recruitment. For adjustment and stratification variables, missing values were assigned to a separate category. Exposure information was either missing or reported as unknown for <5% of women for all potential confounders. Analyses of ovarian cancer risk were restricted to women without bilateral oophorectomy; those of endometrial and cervical cancer were restricted to women without hysterectomy; those of fallopian tube cancer were restricted to women without either bilateral oophorectomy or hysterectomy (as the fallopian tubes might potentially be removed as part of either surgery). Analyses were performed in Stata-14.¹⁶¹ Figures were drawn in R¹⁶² using Matthew Arnold's "Jasper" package.¹⁶³

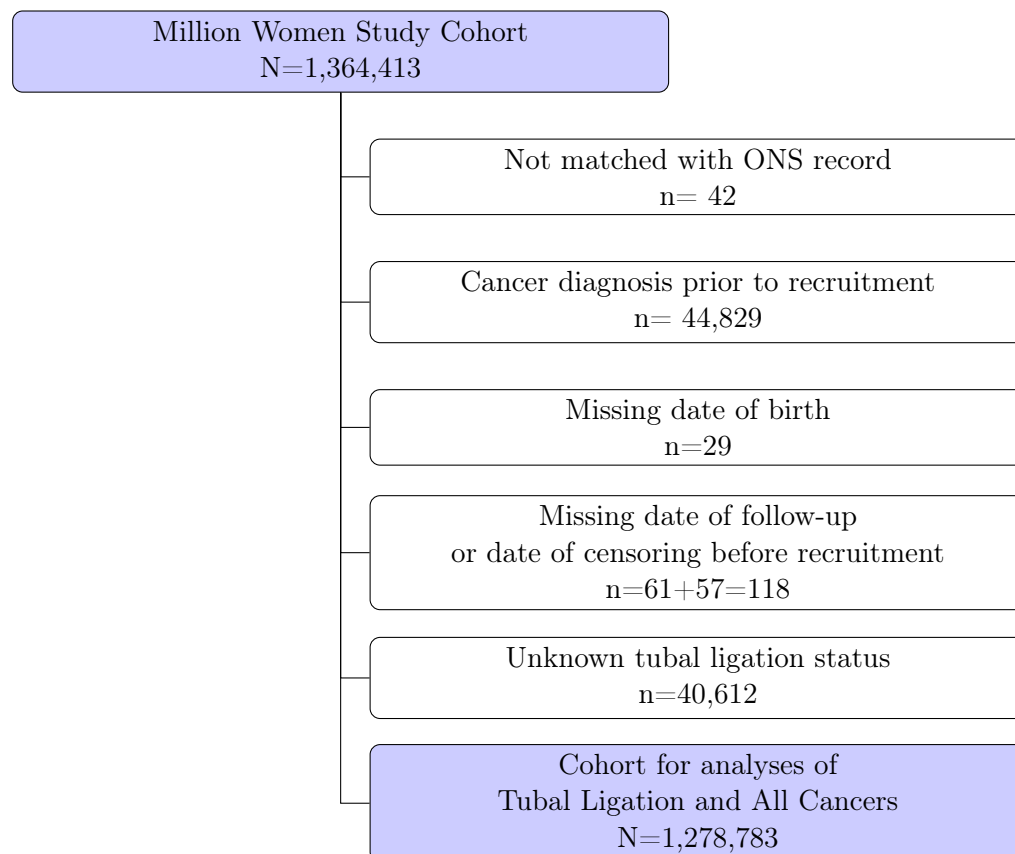


Figure 6.1: *Participants inclusion/exclusion flow diagram for analyses of tubal ligation and all cancers*

6.4 RESULTS

The study population included 1,278,783 women without prior cancer, with a mean age of 56.1 (SD 4.9) at recruitment. 23% of participants (294,724) reported a tubal ligation, at an average age of 34.5 (SD 5.3).

Table 6.1 shows baseline characteristics of the cohort by tubal ligation status. Tubal ligation status was associated with considerable differences in other lifestyle and reproductive factors. Compared to women without a tubal ligation, women who reported a tubal ligation were more likely to be in the lowest quintile of socioeconomic status (25% vs 18%), to have used the oral contraceptive pill (70% vs 57%) or menopausal hormones (58% vs 48%), and to be a current smoker at recruitment (27% vs 19%), and were less likely to be nulliparous (3% vs 13%).

After a mean follow-up of 13.8 years (SD 3.4), a total of 167,430 cancers had been registered. The average age at cancer diagnosis was 65.6 years (SD 6.5). There was no association between tubal ligation and risk of all cancers combined (RR: 1.00, 95% CI: 0.98-1.01).

Figure 6.2 displays adjusted RRs and 95% confidence intervals (CIs) of 26 site- or type-specific cancers amongst women with versus without tubal ligation. As previously seen in Chapter 5, there was a 20% reduction for ovarian cancer risk (RR: 0.80, 95% CI: 0.76-0.85; $p < 0.001$). Interestingly, there was a reduction of similar magnitude for peritoneal cancers (RR: 0.81, 95% CI: 0.66-0.98; $p = 0.03$) (ICD-10 code C48, including retroperitoneal cancers). There was also a large reduction in risk of cancers of the fallopian tube, but confidence intervals were wide (RR: 0.60, 95% CI: 0.37-0.96; $p = 0.04$).

A small increase in risk of lung cancer was noted (RR: 1.09, 95% CI: 1.05-1.13; $p < 0.001$). This may well be due to residual confounding from smoking – as women with tubal ligation were more likely to be current smokers (27% vs. 18%), and to smoke more

cigarettes per day (7% vs 4% smoking 20+/day), than women without tubal ligation. When an interaction term was modelled between smoking and tubal ligation, there was no association between tubal ligation and lung cancer risk amongst never-smokers (RR: 1.04; 95% CI: 0.91-1.19), although there was no evidence that the association between tubal ligation and lung cancer risk varied significantly by smoking status (p-heterogeneity=0.3) (see Figure 6.3). Given the strength of the association between smoking and lung cancer,¹⁰⁴ even a relatively small failure to fully capture smoking history might well lead to substantial residual confounding of an association with lung cancer.

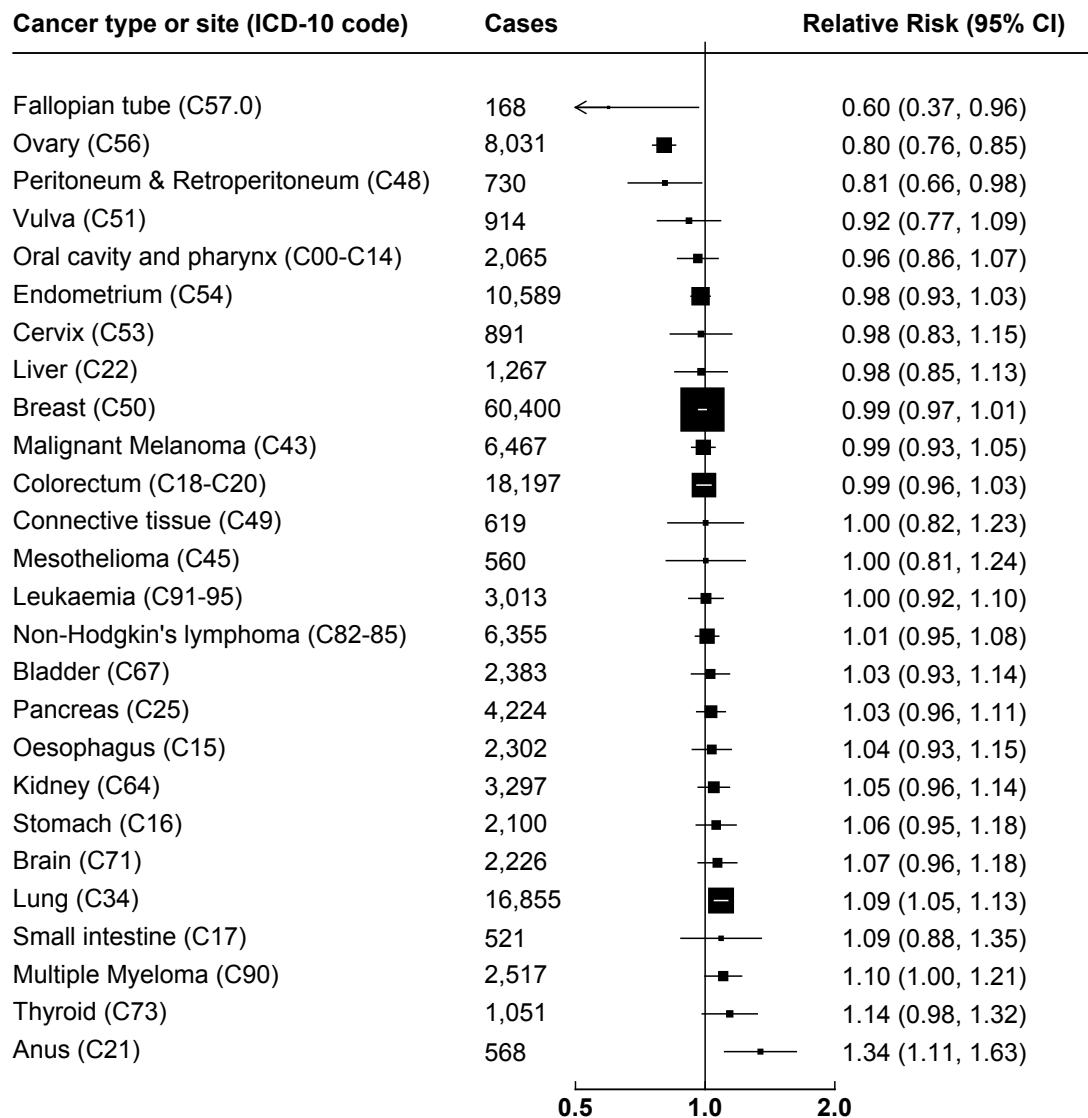
There was an increased risk of anal cancer (RR: 1.34, 95% CI: 1.11-1.63; p=0.002), as has been previously reported in analyses of this cohort.⁵ This association persisted after restricting the analysis to never-smokers (RR: 1.49, 95% CI: 1.08-2.06; p=0.01), and (in a separate analysis) to never-users of the oral contraceptive pill (RR: 1.72, 95%CI: 1.21-2.47; p=0.003), although the reduction in statistical power meant that confidence intervals were wide.

There were no significant associations between tubal ligation and risk of cancers at the other sites, including endometrial (RR: 0.98, 95% CI: 0.93-1.03), cervical (RR: 0.98, 95% CI: 0.83-1.15), breast (RR: 0.99, 95% CI: 0.97-1.01) and colorectal cancers (RR: 0.99, 95% CI: 0.96-1.03).

CHAPTER 6. TUBAL LIGATION AND RISK OF 26 SITE-SPECIFIC CANCERS

Table 6.1: *Characteristics of the study population at recruitment, and details of follow-up, according to tubal ligation status*

	Tubal ligation		
	No	Yes	All women
Characteristics			
Number of women	984,059	294,724	1,278,783
Mean (SD) age at recruitment (years)	56.3 (5.0)	55.4 (4.3)	56.1 (4.9)
Socioeconomic status, lowest quintile, % (n)	18.1 (176,985)	25.1 (73,503)	19.7 (250,488)
Nulliparous, % (n)	13.3 (130,818)	2.6 (7,558)	10.8 (138,376)
Ever use of oral contraceptive pill, % (n)	56.5 (551,799)	69.5 (202,511)	59.5 (754,310)
Ever use of menopausal hormones, % (n)	47.5 (463,250)	58.2 (169,199)	50.0 (632,449)
Hysterectomy, % (n)	21.9 (215,280)	29.1 (84,514)	23.6 (299,794)
Mean (SD) body mass index (kg/m ²)	26.1 (4.6)	26.7 (4.9)	26.2 (4.7)
Current smoker, % (n)	18.5 (171,593)	26.9 (74,647)	20.4 (246,240)
Strenuous exercise ≥ once/week, % (n)	39.5 (375,006)	37.1 (105,032)	38.9 (480,038)
Alcohol intake, ≥ 7 units/week, % (n)	23.3 (227,830)	24.8 (72,471)	23.7 (300,301)
Follow-up for cancer incidence			
Woman-years of follow-up for incidence (100,000s)	136.0	40.5	176.5
Mean follow-up time per woman (SD)	13.8 (3.4)	13.8 (3.3)	13.8 (3.4)
Number of incident cancers	129,531	37,899	167,430
Notes: Means and percentages are calculated excluding missing values for the variable of interest.			
n: number of women. SD: standard deviation.			



Adjusted for age, region, socioeconomic status, age at first birth, parity, hysterectomy, smoking, alcohol intake, exercise, BMI, and use of OCP and HRT.
 Analysis of fallopian tube cancer restricted to women without bilateral oophorectomy or hysterectomy.
 Analysis of ovarian cancer restricted to women without bilateral oophorectomy. Analyses of endometrial and cervical cancers restricted to women without hysterectomy.

Figure 6.2: Relative risk of 26 site- or type-specific cancers in women with versus without a history of tubal ligation

CHAPTER 6. TUBAL LIGATION AND RISK OF 26 SITE-SPECIFIC CANCERS

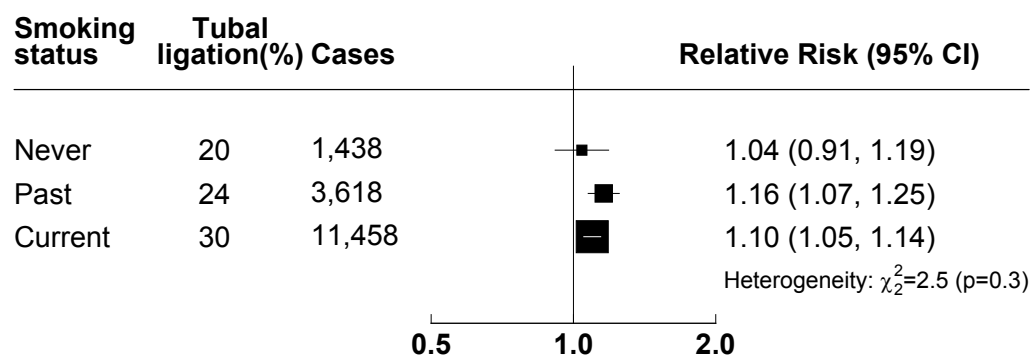


Figure 6.3: Relative risk of lung cancer in women with versus without a history of tubal ligation, by smoking status

6.5 DISCUSSION

This is the largest and most comprehensive analysis to date of the association between tubal ligation and subsequent risk of cancer. The results showed that tubal ligation was significantly associated with reduced risks of ovarian, peritoneal, and fallopian tube cancers and an increased risk of anal cancer, but was not associated with breast, endometrial or cervical cancer, as had been suggested by others.

6.5.1 Tubal ligation and risk of ovarian, peritoneal and fallopian tube cancers

The reduction in ovarian cancer risk has been reported previously, as summarised in several meta-analyses.^{1,127,167} As shown in Chapter 5 (and published¹⁵⁶), there is significant heterogeneity by histological type of ovarian cancer: tubal ligation was associated with a modest reduction of serous tumours overall (RR: 0.84, 95% CI: 0.77-0.92), and about a 20% reduction in risk of high-grade serous carcinoma, but almost a halving of risk of endometrioid and clear cell tumours. These findings are consistent with recent hypotheses regarding different origins of the ovarian cancer histotypes, including the suggestion that many cases of high-grade serous ovarian cancer may arise from precursor lesions within the fallopian tubes, while some endometrioid and clear cell tumours may develop from endometriosis.^{21,22} Although the precise mechanism by which tubal ligation reduces the risk of ovarian cancer is unclear, part of the explanation may well be that it acts as a physical barrier to pro-cancerous substances passing through the fallopian tubes to the ovaries (whether endometriosis, fallopian tube epithelial cells, infectious agents, or exogenous carcinogens).

Few other studies have reported on the association between tubal ligation and risk of peritoneal cancer, and they have tended to be limited by the number of cases. This

CHAPTER 6. TUBAL LIGATION AND RISK OF 26 SITE-SPECIFIC CANCERS

analysis included 730 cases of peritoneal cancer; two small retrospective studies have reported their findings – one study of 62 cases found no association;¹⁸⁵ and another of 22 cases reported a reduced risk.^{117,186}

To my knowledge, this is the first study to report that tubal ligation is associated with a significant reduction in risk of fallopian tube cancer, as few previous studies have had sufficient cases to be able to explore this relatively rare malignancy. A registry-based study in Finland found that tubal ligation was associated with a non-significant reduction in risk of fallopian tube cancer (OR: 0.74, 95% CI: 0.42-1.30), in a multivariate analysis including 189 cases and 1764 controls.¹⁸⁷ An international retrospective study, including 103 cases of fallopian tube cancer and 980 matched controls, also found that tubal ligation was associated with a non-significant reduction in risk of fallopian tube cancer (OR: 0.64, 95% CI: 0.31-1.28).¹⁸⁸

The finding that tubal ligation is associated with a reduction in risk of ovarian, peritoneal and fallopian tube cancers perhaps reflects their histological and clinical similarity, and possible similar site of origin.²³ The majority of the peritoneal and fallopian tube cancers with specified histology were serous carcinomas, which is also the most common histotype of ovarian cancer.¹⁵⁶ There is increasing evidence from histological, molecular and genetics studies that many high-grade serous carcinomas found in the ovary or peritoneum may have originated from precursor lesions in the fallopian tube.^{189,190} While the observation that tubal ligation is associated with a reduction in the risk of all three cancer types (ovarian, peritoneal and fallopian tube) does not prove that many of the cancers at these sites share a common origin in the fallopian tube epithelium, it would be consistent with this hypothesis.

As the majority of the putative precursor lesions for high-grade serous cancers are found within the fimbrial end of the fallopian tube,^{40,50} adjacent to the ovary, it is perhaps surprising that tubal ligation (which normally involves surgery to the mid-portion of the

tube) should reduce the risk of such tumours. It may be that, historically, many surgical sterilization procedures involved cutting and removing a portion of the fallopian tube, rather than simply clipping the tube (as is now more common).

Removal of the fallopian tubes (bilateral salpingectomy) may confer a greater risk reduction of ovarian cancer than tubal ligation alone: one cohort study in Sweden,¹⁹¹ and two case-control studies in Denmark¹⁷⁵ and the USA,¹⁹² all found estimates of effect that suggested a greater risk reduction for bilateral salpingectomy than for simple tubal sterilization. However, for all three studies, confidence intervals were wide and overlapping, so that there was not a significant difference in the associations with the two procedure types. A recent meta-analysis of these studies estimated that bilateral salpingectomy was associated with about a halving of risk of ovarian cancer (OR: 0.51, 95% CI: 0.35-0.75).¹⁹³

One of the limitations of this study is that the available data do not have information on the type of surgical procedure performed, which makes it impossible to explore possible differences in association between different surgical techniques. Although MWS is linked to electronic hospital records, which provide surgical procedure codes, we only have hospital records for England from April 1997. Given that the median year of tubal ligation surgery in the MWS cohort is 1978 (with 99% of procedures occurring by 1992), data that only start in 1997 would not capture these procedures.

Alternatively, it may be that surgery causes local vascular changes with consequent effects on the tubal epithelium. Unfortunately, it would be very difficult (if not impossible) to investigate this hypothesis using routine clinical samples. One might posit an experiment in which one examines the fallopian tubes of women who have undergone tubal ligation, and then subsequently had the remainder of their fallopian tubes removed (e.g. as part of a hysterectomy and bilateral salpingoophorectomy), compared to women whose fallopian tubes were removed without prior tubal ligation. However, while such

specimens may exist, it would be difficult to ensure an unbiased comparison - as the reason for the subsequent removal of the tubes is likely to be serious disease (e.g. actual or suspected cancer). Any such serious disease might well be a more plausible explanation of any changes one found in the fallopian tubes, rather than a post-ligation effect.

6.5.2 Tubal ligation and risk of other cancers

For endometrial cancer, a reduced risk associated with tubal ligation was reported in a Danish cohort,¹⁸² but we and others have not replicated this association.^{194,195} One group of investigators reported that tubal ligation was associated with an increased risk of cervical cancer,¹⁸³ but again, neither other findings nor ours replicated this.¹⁹⁶

Reports of the relationship between tubal ligation and breast cancer risk have been varied, including suggestions of either a reduced^{181,197} or increased risk,¹⁸⁰ although the majority of studies reported no significant association,^{184,198–203} consistent with our findings.

The risk factors for anal cancer in the Million Women Study have previously been reported on in detail;⁵ factors associated with an increased risk of anal cancer included tubal ligation, nulliparity, smoking, use of the oral contraceptive pill, and not living with a partner. The strongest association was with a history of cervical intraepithelial neoplasia grade 3 (RR: 4.03, 95% CI: 2.59-6.28), reflecting the importance of high-risk strains of human papilloma virus (HPV) as causative agents of both cervical,²⁰⁴ and anal²⁰⁵ cancers.

It is unclear why tubal ligation should be associated with an increased risk of anal cancer, and it may be that this apparent association reflects residual confounding. Our observation that the association persisted even when analyses were restricted to never-smokers, or never-users of the oral contraceptive pill, indicates that these factors (al-

though both associated with tubal ligation and risk of anal cancer) are unlikely to be responsible for the association between tubal ligation and anal cancer risk. Given the importance of the human papilloma virus as a causative agent of anal cancer, one possible alternative explanation would be confounding by sexual behaviour and exposure to HPV. For example, as tubal ligation is usually performed for permanent contraception, it is likely to be more common amongst women who are sexually active, and who thus may be more likely to be exposed to sexually-transmitted infections, such as HPV, and the consequent increased risk of anal cancer. However, we did not find that tubal ligation was associated with other HPV-driven cancers (e.g. cervical cancer). As the Million Women Study questionnaire did not ask for information on sexual activity or sexual history, it was not possible to explore this further.

There have been reports on possible relationships between tubal ligation and other cancers, including possible reductions in risk of colorectal^{206,207} and stomach cancer,¹⁸⁴ and increased risks of thyroid cancer²⁰⁸ and lymphatic and haematopoietic malignancies.¹⁸² This analysis in the Million Women Study did not replicate any of these findings.

The observation that tubal ligation is strongly associated with reduced risks of ovarian, peritoneal and fallopian tube cancers, but is not associated with reduced risks of other hormonally-driven cancers (such as breast or endometrial cancers) might help to point to a possible mechanism of action (assuming the association to be causal). If tubal ligation acted to reduce the risk of ovarian cancer primarily by impairing the blood supply to the ovary, and so reducing ovarian hormone levels (as has been suggested by some authors²⁰⁹⁻²¹¹), then one might expect tubal ligation also to reduce the risks of the other hormonally-driven cancers. The fact that it does not (at least in these data) suggests that an alternative mechanism is at work. For example, tubal ligation may act as a barrier to pro-cancerous agents (whether fallopian tube epithelial cells, endometrial cells - or potentially even exogenous agents or microbes) passing through the fallopian

tubes to reach the ovary and peritoneal cavity.

6.5.3 Strengths and weaknesses of this study

As discussed previously, the strengths of the Million Women Study include its large size (limiting random error due to chance), and its prospective design (limiting systematic error due to bias), together with its detailed information on many factors, allowing for statistical adjustment for many potential confounding variables. The prospective design, with outcome data including incident cancers of any type, is also essential for this type of comprehensive analysis, exploring associations between a single exposure and multiple different outcomes, using standardised analysis methods.

In the case of analyses of tubal ligation and cancer risk, the Million Women Study has the added strength that tubal ligation was a relatively common procedure in this cohort (23% of women reported having had a tubal ligation). By comparison, a nationwide register-based study in Denmark was relatively under-powered due to the rarity of tubal ligation in that population (around 3%).¹⁷⁵

As previously mentioned, one of the weaknesses of the study is the absence of detailed information on the type of surgical procedure performed. This precludes any investigation of whether partial or complete resection of the fallopian tube (salpingectomy) is associated with a greater reduction in risk than simple occlusion (e.g. with clips). This would be of interest, as there have been several studies suggesting that bilateral salpingectomy may be associated with a reduced risk of ovarian cancer (as discussed above).

6.5.4 Conclusion

The large size of the cohort, the individual information on possible confounding factors, and the complete and long follow-up, provided reliable estimates of risks associated with tubal ligation for 26 specific cancer sites, even relatively uncommon ones. No association was found between tubal ligation and the risk of cancers of the endometrium, breast, or cervix. By contrast, tubal ligation is associated with a clear reduction in risk of ovarian cancer, a reduction of similar magnitude of peritoneal cancer, and a reduction of fallopian tube cancer. That tubal ligation is associated with a reduced risk of cancers of the ovary, peritoneum and fallopian tube, but not of other hormonally-related cancers, is consistent with the hypothesis that many of the cancers at these sites have a shared origin in the fallopian tube, and that tubal ligation reduces cancer risk by acting as a barrier to cells, carcinogens or other agents reaching the ovary and peritoneal cavity, rather than by affecting hormone levels.

CHAPTER 7

REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

7.1 INTRODUCTION

Many epidemiological studies have reported a reduced risk of ovarian cancer amongst women who have had children (parous), finding decreasing risk with increasing parity (number of births).^{20,107–122}

Long durations of breastfeeding have also been associated with a reduced risk of ovarian cancer,^{123–125} although there is considerable heterogeneity in the findings, not all of which adjusted adequately for parity, which is highly correlated with total duration of breastfeeding. The World Cancer Research Fund's Continuous Update Project has concluded that "The evidence suggesting that breastfeeding protects against ovarian cancer is limited."¹²⁶

The majority of cases of ovarian cancer are epithelial tumours, of which the most common histological type is high-grade serous carcinoma (about 60%); the next most common histotypes are endometrioid, clear cell, mucinous, and low-grade serous carci-

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

nomas, each of which accounts for less than 20% of the total.²¹² Recent evidence from histopathological and genetic studies suggests that the different histotypes of ovarian cancer may have distinct, possibly extra-ovarian, origins. For example, it is suggested that many high-grade serous tumours may arise from precursor lesions within the fallopian tubes, while some endometrioid and clear cell tumours may develop from endometriosis²¹ (discussed in more detail in Chapter 2). Many factors, including use of oral contraceptives,¹⁹ use of menopausal hormones,¹⁸ and smoking¹⁰³ show variation in risk by ovarian tumour histotype (see Chapter 3).

Some^{108,112,119,122} but not all^{20,111,113} epidemiological studies have reported that some associations between childbearing patterns and ovarian cancer varied by histotype, although few formally tested for heterogeneity in this and there is little consistency in the findings across studies.^{20,108,111–113,119,122} No single previous study included more than 2,000 cases of ovarian cancer, too few to examine reliably for heterogeneity by histotype (a record-linkage Swedish study was larger but had extremely limited information on potential confounders - only age at diagnosis and age at first birth were adjusted for¹⁰⁷).

The largest pooled analysis of retrospective studies included 5,566 cases of ovarian cancer, and results on parity and risk of ovarian cancer histotypes were reported just as part of an analysis of a genetic risk score in relation to environmental factors.²⁶ A recent pooled analysis of 21 prospective studies from the Ovarian Cancer Cohort Consortium, including 5,584 cases of invasive epithelial ovarian cancer, found strong evidence of an association between parity and ovarian cancer risk, with heterogeneity by histotype, but no significant association with breastfeeding, either overall or by histotype.²⁷

Large numbers are needed to examine reliably the association between reproductive factors (parity and breastfeeding) and risk by ovarian cancer histotype, and so this chapter reports findings from a prospective study of one million UK women with over 8,700 incident cases of ovarian cancer.

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

7.2 AIM

This chapter investigates the association between childbearing patterns (parity and breast-feeding) and risk of ovarian cancer histotypes. The main aim of these analyses was to look separately at nulliparity compared to women with one child (para-1), and then by number of births amongst parous women.

7.3 METHODS

The methods of recruitment and follow-up for the Million Women Study have been described in detail in Chapter 4, and are summarised below.

7.3.1 Study design, data collection and follow-up

The Million Women Study is a cohort of 1.3 million UK women, aged 50-64 at the time of recruitment (1996-2001). Participants were invited to take part in the study via letters sent with invitations to attend mammographic screening for breast cancer. The women completed questionnaires at recruitment, including information on reproductive, sociodemographic, and health factors. The study methods have been described in detail elsewhere,^{147,148} and study questionnaires can be viewed at <http://www.millionwomenstudy.org>.

Follow-up was via record linkage to the National Health Service Central Register (NHSCR), based on unique NHS number and other personal identifiers: all participants were ‘flagged’, so that the study investigators were routinely notified of cancer registrations or deaths. Cancer diagnoses were coded by NHSCR using the 10th revision of the WHO International Classification of Diseases (ICD-10),¹⁵¹ with morphology codes from the 2nd and 3rd edition of the International Classification of Diseases for Oncology

(ICD-O).^{154,155}

All participants gave written consent to follow-up at recruitment. The study received ethical approval from the Oxford and Anglia Multi-Centre Research Ethics Committee (MREC 97/01).

7.3.2 Exposure variables

Information on reproductive factors was taken from the recruitment questionnaire. Parity was defined by the number of full-term pregnancies, as reported in response to the question: "How many children have you had? (please include stillbirths; it is not necessary to include miscarriages)". Women were also asked to record the date of birth of each child, which was used for analyses of the woman's age at first or last birth.

Total duration of breastfeeding was calculated as the sum of the duration of breastfeeding for all children, as reported in response to the question: "for how many months did you breastfeed each child, if at all?" [Note that the question on breastfeeding was only added to the questionnaire after about 10% of the cohort had been recruited, and thus information is missing for this first batch of women].

7.3.3 Outcome

The outcome of interest was ovarian cancer (ICD-10 C56). For analyses of histotype, the outcome was split into the four most common specific histological groups: serous (ICD-O codes 8441-8442, 8451, 8460-8463, 9014), mucinous (ICD-O codes 8470-8490), endometrioid (ICD-O codes 8380-8381, 8560, 8570), and clear cell (ICD-O codes 8310, 8313); the remainder of cases were classified as other/ unspecified.

For some analyses, serous and mucinous tumours were further subdivided, as previously described¹⁵⁶ (also see Chapter 5). Serous tumours were divided into low-grade

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

tumours (defined here as borderline serous tumours (ICD-O codes 8442, 8451, 8462, 8463) or grade 1 serous carcinomas), and high-grade serous carcinomas (defined here as grade ≥ 2 serous carcinomas). Mucinous tumours were split into mucinous borderline tumours (ICD-O codes 8472-8473) and mucinous carcinomas (ICD-O codes 8470-8471 and 8480-8490); the mucinous carcinomas were not further subdivided by tumour grade, as numbers were small.

7.3.4 Statistical analyses

As part of standard exclusions, women were excluded if they had been diagnosed with any invasive cancer (other than non-melanoma skin cancer, ICD-10 C44) prior to recruitment ($n=44,775$), if they had missing information on follow-up date ($n=81$) or if their date of censoring at ONS was prior to recruitment ($n=60$). For this analysis, women were also excluded if they reported a previous bilateral oophorectomy (or were unsure, or the information was missing) ($n=170,772$); or if they had missing information on parity ($n=3,883$) (see Figure 7.1).

The remaining women ($N=1,144,762$) contributed person-years from the date of recruitment into the study until the date of registration for ovarian cancer, the date of death, or last date of follow-up (31st December 2014) – whichever was soonest. Women were censored at diagnosis of any cancer (other than non-melanoma skin cancer). About 1% of participants were lost to follow-up by the end of 2014 and such women are censored at the date when they were lost, contributing person-years until then.

Cox proportional hazards models were used to estimate hazard ratios (referred to as relative risks (RRs)) of ovarian cancer according to parity, age at first or last birth, and breastfeeding. Attained age was the underlying time variable. There was no evidence of significant violation of the proportional hazards assumption, as assessed by graphical methods and tests based on Schoenfeld residuals (data not shown).

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

Analyses examined associations with parity, either as a binary variable (nulliparous compared to parous), or as an ordered categorical variable (1, 2, 3, 4, 5+) in parous women. [Note that for the purpose of plotting graphs, some analyses were also run with parity grouped as 1, 2, 3, 4+.] A piecewise regression model was used to test for heterogeneity in the magnitude of the reduction in risk between para-1 versus nulliparous compared to the RR per birth for subsequent births, and allowed a single model to be used to examine these two associations simultaneously, whilst allowing the coefficients for the two associations to be different. Analyses of the relative risk per additional birth were restricted to parous women within the piecewise regression model, and parity was treated as a continuous variable, with groups coded as the mean number of births per group (1, 2, 3, 4, 5.5). Where parity was examined as a categorical variable separately by histological type, it was analysed in fewer groups (up to 3+ only), due to small numbers of cases of less-common subtypes in higher parity groups.

Analyses of the timings of births were restricted to parous women with non-missing information on age at first and last births, and were stratified by parity (1, 2, 3+). Both age at first birth, and age at last birth, were treated as ordered categorical variables (age at first birth: <20, 20-24, 25+; age at last birth: <25, 25-29, 30+). Tests for trend used the categorical variables coded with the median age in each age-group (age at first birth: 18, 22, 27; age at last birth: 23, 27, 33). Separate models were run for age at first birth and age at last birth.

Analyses of breastfeeding used the total duration of breastfeeding (when summed for all births), stratified by the number of births (1, 2, 3, 4, 5+), as an ordered categorical variable (none, ≤ 1 month, >1 & <6 months, ≥ 6 & <12 months, 12+ months). Tests for trend used the categorical variable coded by the median duration of breastfeeding in each group (0, 1, 3, 8, 17 months). For analyses of breastfeeding, women were excluded if they had missing information on the duration of breastfeeding (information was missing for

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

approximately 20% of women, as this question was not included in the earliest batch of questionnaires). All breastfeeding analyses were restricted to parous women, and stratified by parity (1, 2, 3, 4, 5+).

All analyses were stratified by geographical region (10 regions corresponding to the areas covered by the cancer registries), and adjusted for use of the oral contraceptive pill (never, ever), tubal ligation (no, yes), family history of breast cancer (no, yes), hysterectomy (no, yes), use of menopausal hormones (never, ever), body mass index (BMI) (<25 kg/m², 25-29 kg/m², 30+ kg/m²), smoking history (never, past, current), and quintiles of socioeconomic status based on the Townsend deprivation index.¹⁵⁷ For adjustment variables, missing values were assigned to a separate category. All information on adjustment variables was taken from the recruitment questionnaire. Exposure information was either missing or reported as unknown for $<5\%$ of women for all potential confounders.

Other factors (including alcohol consumption, physical activity, and age at menarche) were explored as potential confounders, but were not included in the final model as their inclusion made no appreciable difference to the main estimate of effect.

Tests of heterogeneity in the relationship between ovarian cancer and parity or breastfeeding by histotype were performed using a competing risks approach.²

For analyses where comparisons were made between more than two exposure categories, relative risks are presented in figures with group-specific confidence intervals for the log risk in each group (allowing comparisons to be made between any two categories, even if neither is the reference group).^{159,160} Conventional 95% confidence intervals are given in the text. All analyses were performed in Stata-14.¹⁶¹ Figures were drawn in R,¹⁶² using the standard package 'ggplot', or Matt Arnold's "Jasper" package.¹⁶³

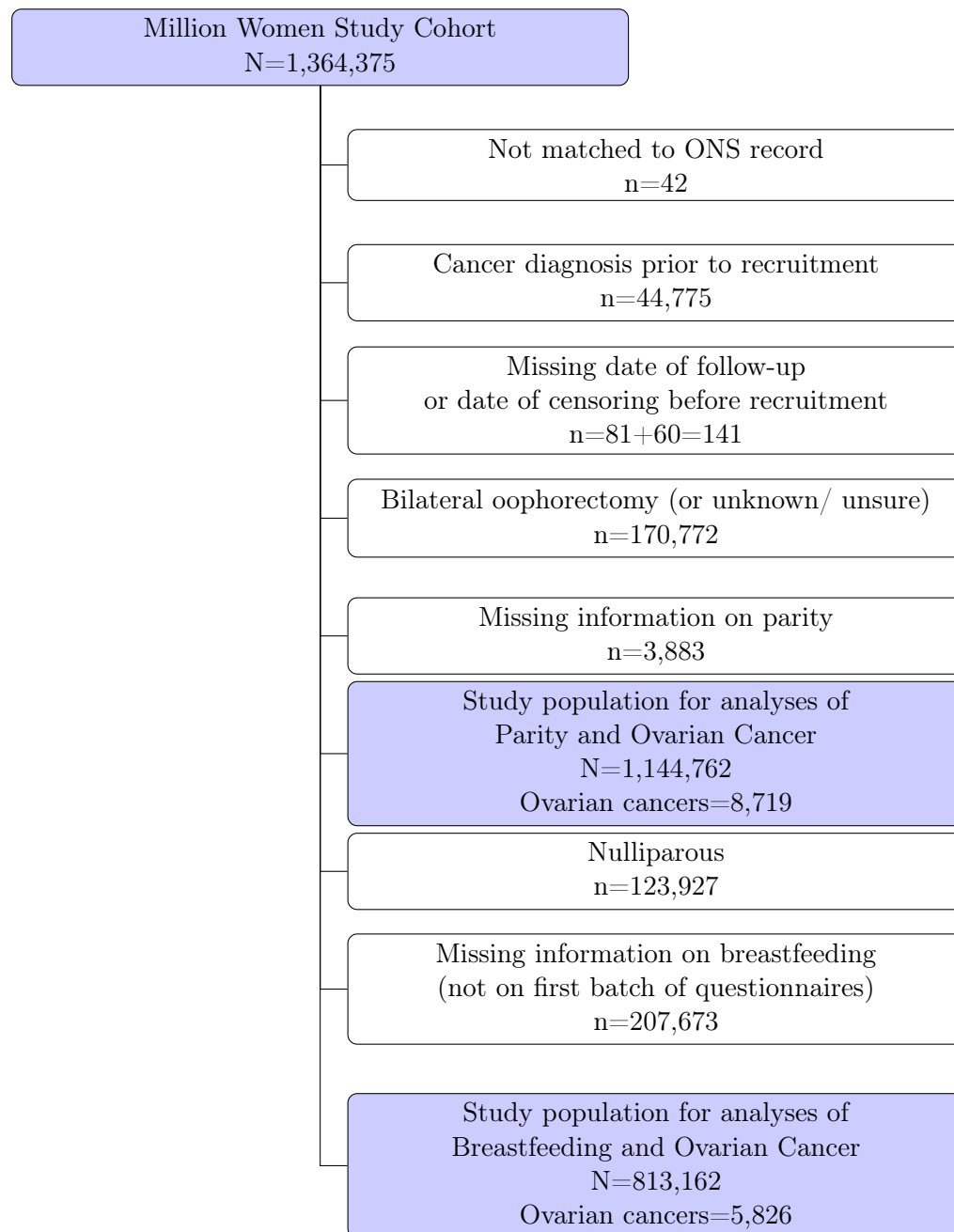


Figure 7.1: Participants inclusion/exclusion flow diagram for analyses of reproductive factors and risk of ovarian cancer

7.4 RESULTS

A total of 1,144,762 women were included in this analysis, with a mean age at recruitment of 56.1 years (standard deviation (SD) 4.8). Overall, 89% of participants reported at least one full-term pregnancy. Table 7.1 shows the characteristics of the study population, by parity. Parous women were more likely to have used the oral contraceptive pill (62% vs 46%) and menopausal hormones (48% vs 42%), to have undergone tubal ligation (24% vs 5%) and hysterectomy (17% vs 10%), and to be current smokers (21% vs 17%), compared to nulliparous women.

During 16.7 million person-years of follow-up (an average of 14.6 (SD 3.7) years per woman), 8,719 incident ovarian cancers were reported, with a mean age at diagnosis of 65.6 years (SD 6.5). Of the 5,848 epithelial cancers with the four most common tumour types, 67% were serous (n=3,916), 15% mucinous (n=896), 11% endometrioid (n=622) and 7% clear cell (n=414). Of the remaining cases (n=2,871), the majority were recorded as unspecified carcinoma or adenocarcinoma.

7.4.1 Parity and risk of ovarian cancer

Overall, parous women had an estimated 26% lower risk of ovarian cancer than nulliparous women (RR: 0.74, 95% CI: 0.70-0.79), after adjustment for age, region, socioeconomic status, tubal ligation, family history of breast cancer, hysterectomy, BMI, smoking, and use of contraceptive or menopausal hormones.

The reduction in risk of ovarian cancer was greatest for the first birth, with almost a 20% reduction in risk compared to nulliparous women (RR: 0.82, 95% CI: 0.76-0.89); there were further reductions in risk with subsequent births, but of smaller magnitude (<10% per birth) (see Figure 7.2).

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

Given this marked difference in the reductions in risk associated with the first versus subsequent births, it was not reasonable to model parity simply as a relative risk per birth (likelihood ratio test for model with parity modelled as a categorical variable versus continuous: $p=0.002$). Instead, a piecewise regression model was used, with women with one child (para-1) as the baseline; this allowed for distinct relative risks to be modelled for nulliparous women compared to para-1 women, and for additional births amongst parous women.

In the piecewise regression model, nulliparous women had a 24% higher risk of ovarian cancer than women with one child (RR: 1.24, 95% CI: 1.16-1.33), and there was significant heterogeneity by histotype (heterogeneity: $p=0.01$). Nulliparous women had about a 50% higher risk of endometrioid (RR: 1.49, 95% CI: 1.18-1.89, $p=0.001$) and almost a 70% higher risk of clear cell tumours (RR: 1.68, 95% CI: 1.29-2.20, $p<0.001$), with a smaller increase in risk of mucinous tumours (RR: 1.26, 95% CI: 1.02-1.57, $p=0.03$), and a small and non-significant increase in risk of serous tumours, the most common histotype (RR: 1.10, 95% CI: 0.99-1.22) (see Figure 7.3 (A)). There was no significant heterogeneity in the association between low-grade and high-grade serous tumours (heterogeneity: $p=0.5$), or between mucinous borderline tumours and mucinous carcinoma (heterogeneity: $p=0.3$).

Among parous women, each additional birth was associated on average with a 6% reduction in risk of ovarian cancer (RR: 0.94, 95% CI: 0.92-0.96). This association was heterogeneous by histotype (heterogeneity: $p=0.0006$), with a 25% reduction in risk per birth of clear cell tumours, (RR: 0.75, 95% CI 0.65-0.85, $p<0.001$), a lesser reduction in the risk of endometrioid tumours (RR: 0.88, 95% CI: 0.80-0.97, $p=0.009$), a relatively small reduction in the risk of serous tumours (RR: 0.93, 95% CI: 0.90-0.96, $p<0.001$), and no significant reduction in risk of mucinous tumours (RR: 0.95, 95% CI: 0.89-1.02, $p=0.2$) (see Figure 7.3 (B)). When serous tumours were divided, there was a significant reduction

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

in risk of serous borderline and low-grade carcinoma (RR: 0.84, 95% CI: 0.76-0.93), but not of high-grade serous carcinoma (RR: 0.97, 95% CI: 0.92-1.02) (heterogeneity: $p=0.01$). There was no significant heterogeneity in the association between mucinous borderline tumours and mucinous carcinoma (heterogeneity: $p=0.4$).

The 24% increase in the risk of ovarian cancer amongst nulliparous women compared to women with one child was significantly greater than would have been expected based on the 6% reduction in risk per birth found amongst parous women (heterogeneity from a piecewise regression model: $p=0.0003$). The association between parity and risk of ovarian cancer did not vary significantly by subgroups (see Figure 7.5). There was no significant trend seen with increasing age at first or last birth, either for ovarian cancer overall, or for the four main histotypes (see Figure 7.4).

7.4.2 Breastfeeding and risk of ovarian cancer

Amongst parous women with information on breastfeeding ($N=813,162$), 68% (554,695) reported ever breastfeeding, with a mean duration of 8.1 (SD 9.4) months overall, or an average of 3.3 (SD 3.4) months per child. Women who reported breastfeeding (of any duration) had a small, borderline significant, reduction in the risk of ovarian cancer overall, compared to women who never breastfed (RR: 0.94, 95% CI: 0.89-1.00, $p=0.04$).

However, there was evidence of a significantly-reduced risk of ovarian cancer amongst women who reported an average duration of breastfeeding of ≥ 6 months (RR: 0.90, 95% CI: 0.83-0.98 for duration ≥ 6 & < 12 months; RR: 0.86, 95% CI: 0.79-0.94 for duration 12+ months) (see Figure 7.6). Overall, there was about a 10% reduction in the relative risk of ovarian cancer per 12-month increase in duration of breastfeeding (RR: 0.89, 95% CI: 0.84-0.94, $p<0.001$), with no significant heterogeneity by histotype (heterogeneity: $p=0.3$) (see Figure 7.7).

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

Table 7.1: *Characteristics of the study population at recruitment, and details of follow-up for ovarian cancer, according to parity*

	Parity		All women
	Nulliparous	Parous	
Characteristics			
Number of women	123,927	1,020,835	1,144,762
Mean (SD) age at recruitment (years)	56.1 (5.0)	56.1 (4.8)	56.1 (4.8)
Socioeconomic status, lower third, % (n)	30.7 (37,774)	32.3 (327,624)	32.2 (365,398)
Mean (SD) age at menarche	12.9 (1.6)	13.0 (1.6)	13.0 (1.6)
Ever use of oral contraceptive pill, % (n)	45.6 (56,054)	61.6 (624,076)	59.9 (680,130)
Tubal ligation, % (n)	5.0 (6,149)	23.8 (239,239)	21.7 (245,388)
Ever use of menopausal hormones, % (n)	42.0 (51,683)	47.5 (479,976)	46.9 (531,659)
Hysterectomy, % (n)	10.3 (12,796)	16.9 (172,168)	16.2 (184,964)
Mean (SD) age at natural menopause	48.7 (4.4)	49.2 (4.2)	49.1 (4.2)
Family history of breast cancer, % (n)	10.0 (11,615)	9.8 (94,349)	9.9 (105,964)
Mean (SD) body mass index (kg/m ²)	25.7 (4.8)	26.2 (4.6)	26.2 (4.7)
Current smoker, % (n)	17.1 (20,130)	20.5 (197,543)	20.2 (217,673)
Strenuous exercise ≥once/week, % (n)	39.9 (47,959)	39.2 (385,852)	39.3 (433,811)
Alcohol intake, ≥7 units/week, % (n)	27.2 (33,473)	23.6 (238,840)	24.0 (272,313)
Follow-up for ovarian cancer incidence			
Woman-years of follow-up for incidence (100,000s)	17.8	149.5	167.3
Mean follow-up time per woman (SD)	14.3 (4.0)	14.6 (3.6)	14.6 (3.7)
Number of incident cancers	1,266	7,453	8,719
Notes: Means and percentages are calculated excluding missing values for the variable of interest. n: number of women. SD: standard deviation.			

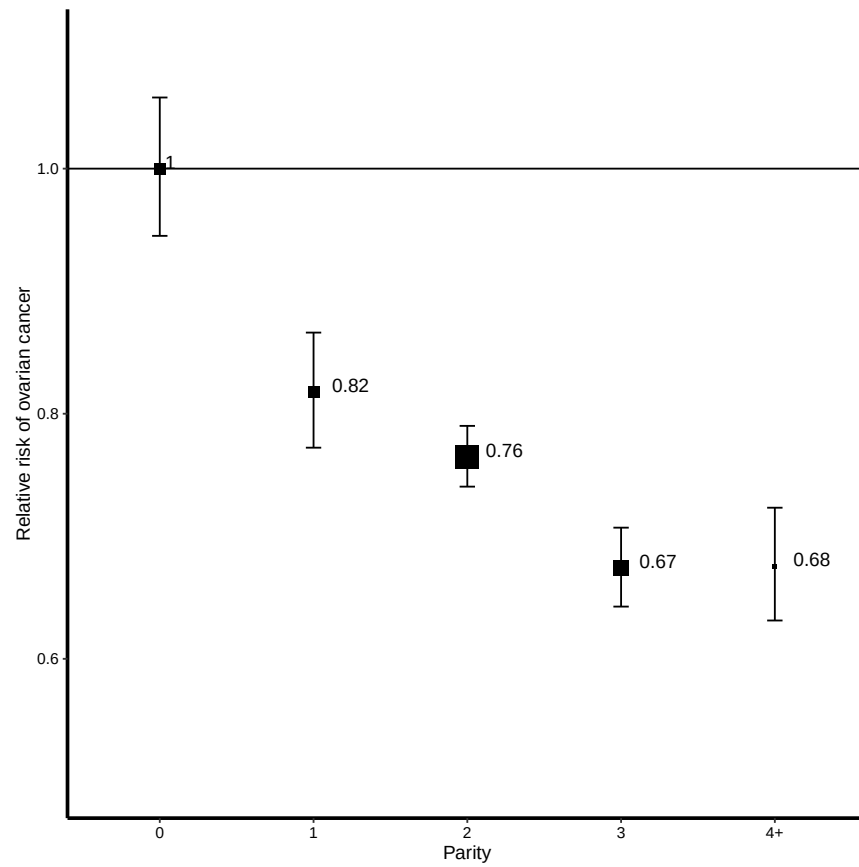


Figure 7.2: *Relative risk of ovarian cancer by parity*

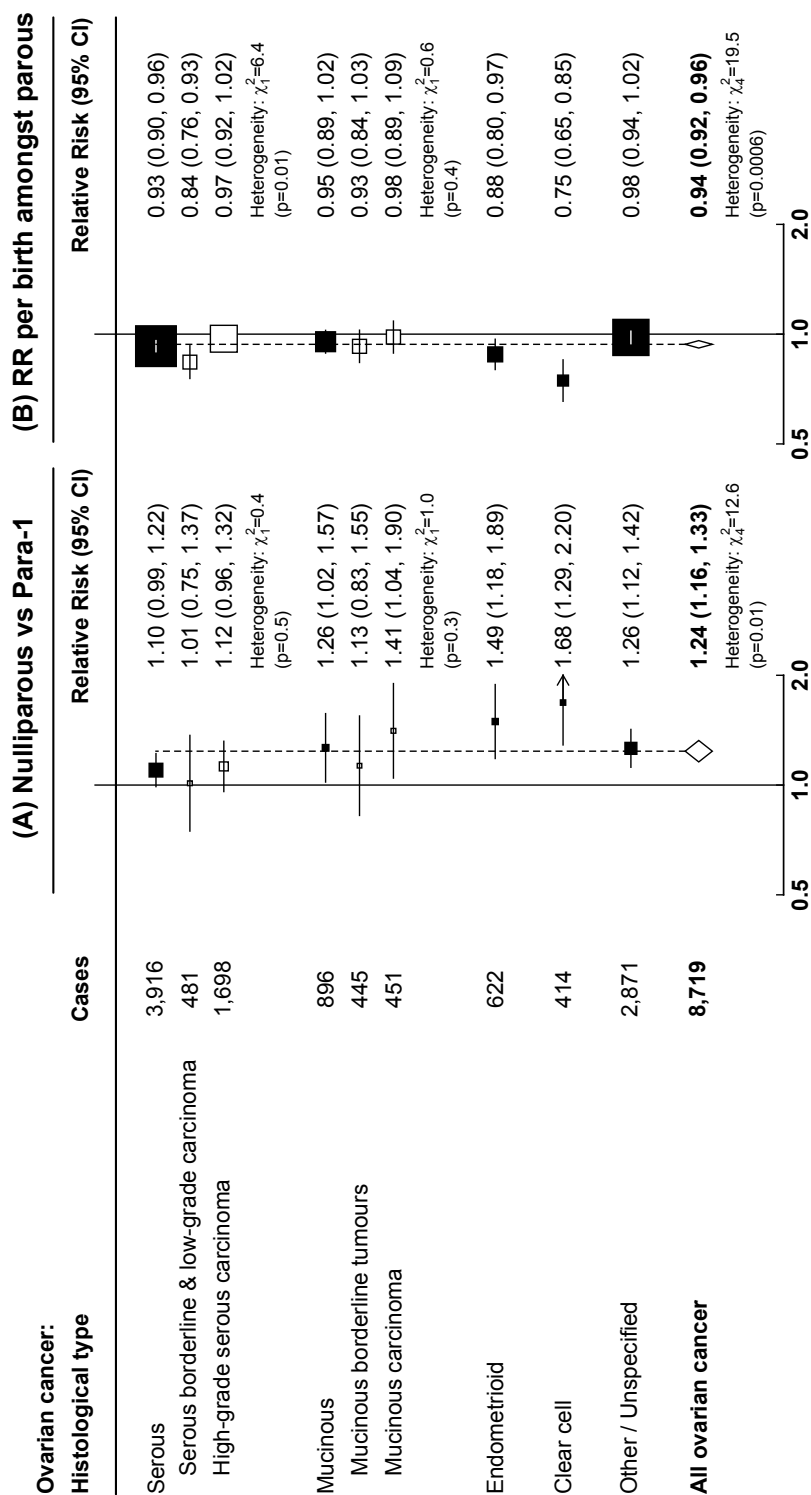


Figure 7.3: Relative risk of histological types of ovarian cancer by parity

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

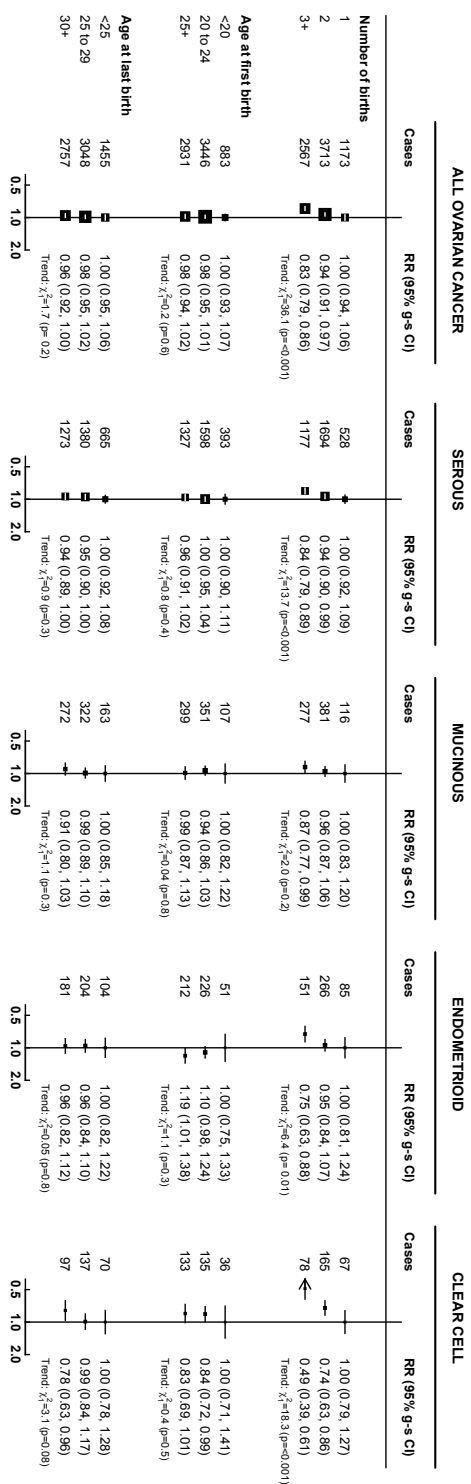


Figure 7.4: Relative risk of histological types of ovarian cancer by number of births and age at first and last births

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

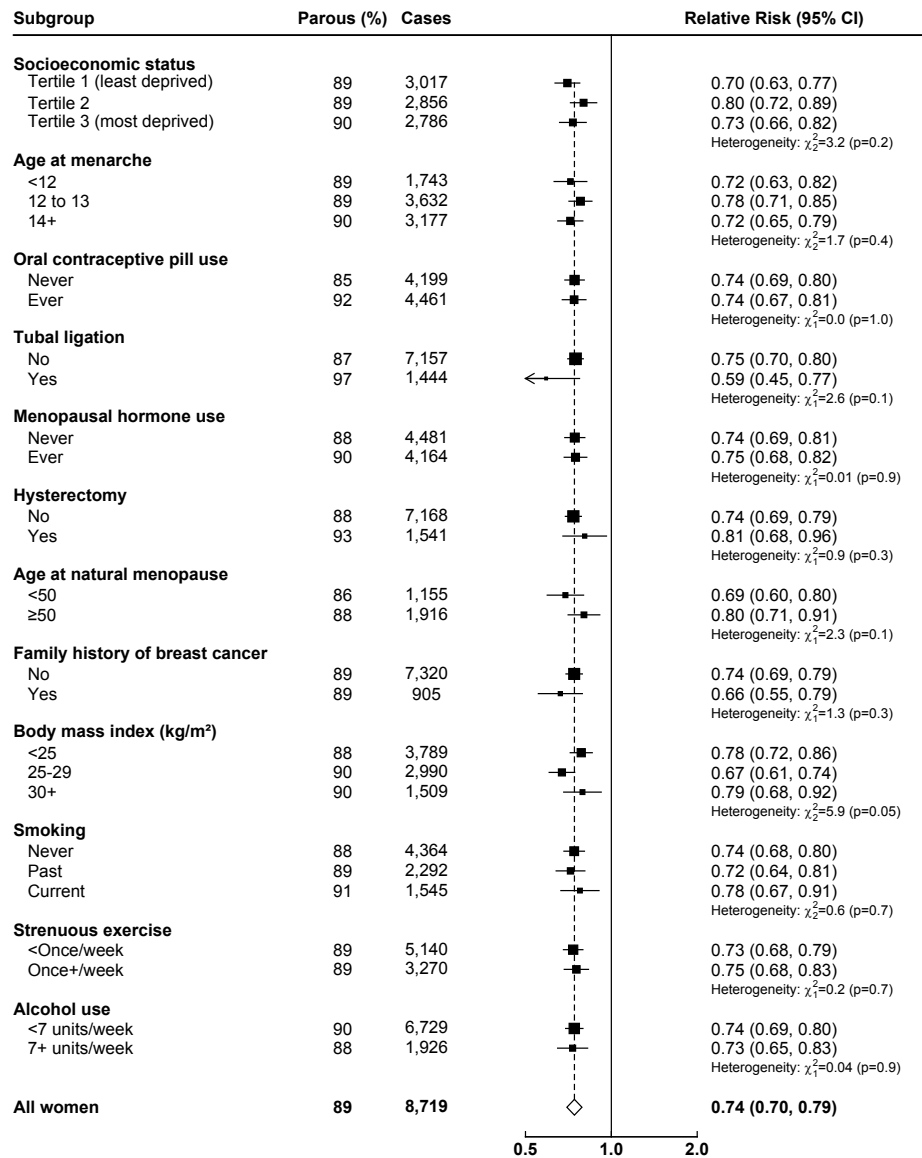


Figure 7.5: Relative risk of ovarian cancer amongst parous vs nulliparous women, by subgroups

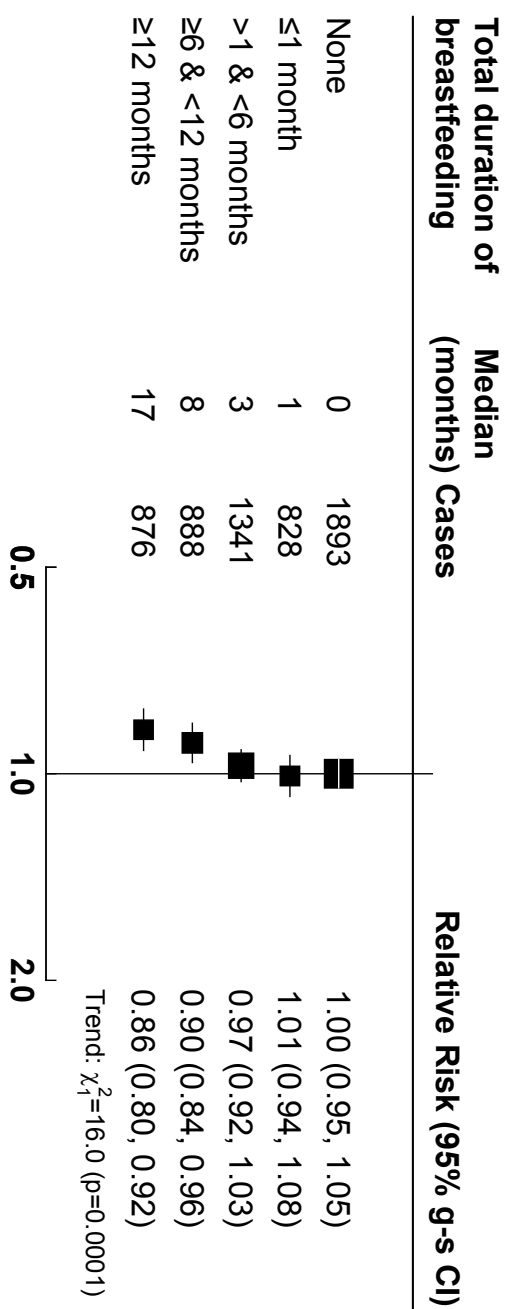


Figure 7.6: *Relative risk of ovarian cancer with increasing duration of breastfeeding*

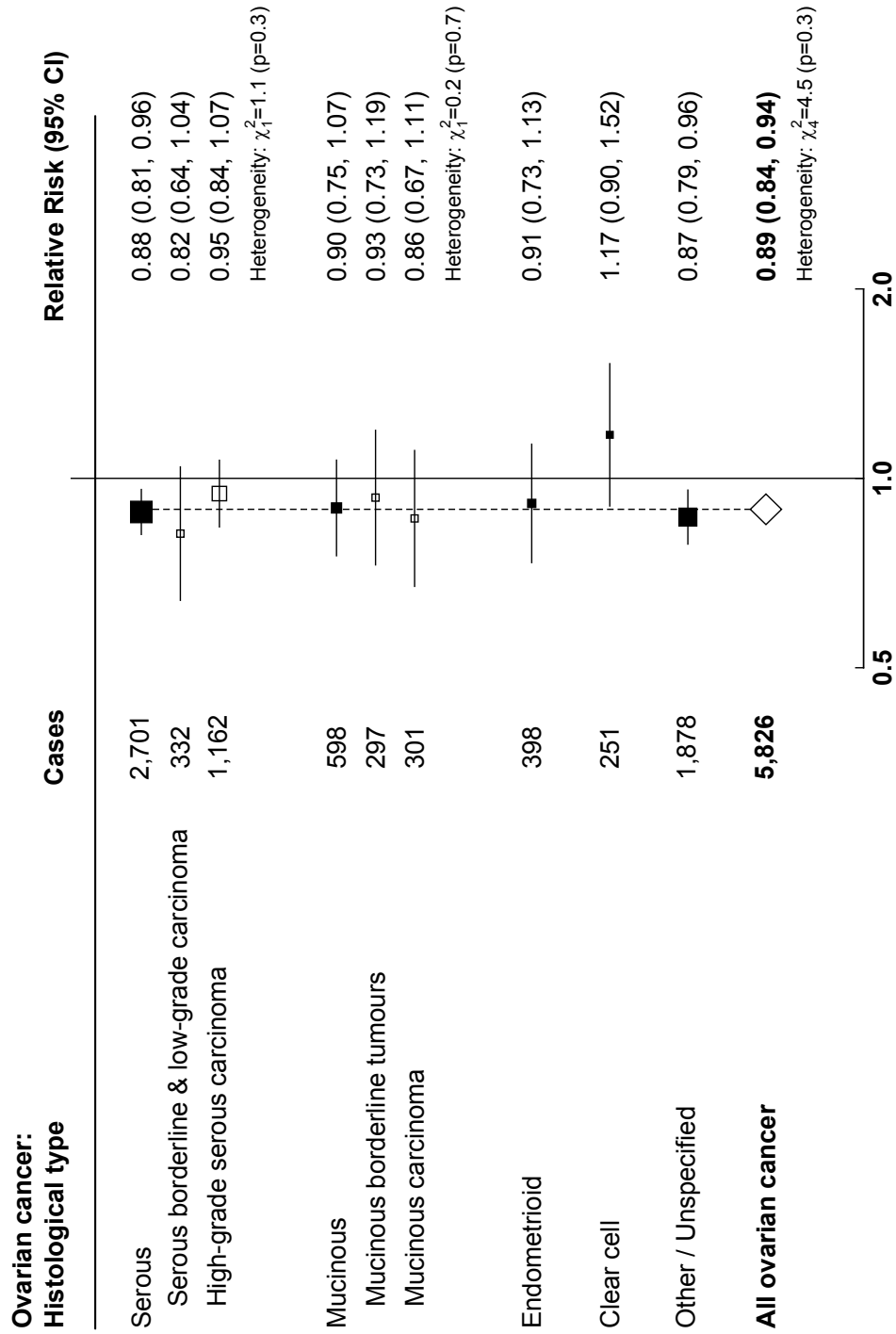


Figure 7.7: Relative risk of ovarian cancer histotypes per 12-months breastfeeding

7.5 DISCUSSION

These analyses have shown, using data from a large prospective study with 8,719 incident cases of ovarian cancer, that there is strong evidence of a (small) reduction in ovarian cancer risk with additional births amongst parous women, in agreement with other studies.

However, there was almost a 25% greater risk of ovarian cancer amongst nulliparous women compared to women with just one child, which was substantially greater than the 6% reduction in risk seen with each birth in parous women. Such a difference has been noted previously,²⁰ but had rarely been formally tested before. For both associations, there were differences by histotype, but the pattern of associations by histotype differed for these two measures of parity (i.e. nulliparous vs para-1, compared to RR-per-birth amongst parous).

The reduction in risk of ovarian cancer seen with parity varied significantly between the histological types. The greatest differences were seen for endometrioid and clear cell tumours: each additional birth amongst parous women was associated with about a 25% reduction in risk of clear cell tumours, and about a 10% reduction in risk of endometrioid tumours, and there was about a 50-70% relative excess risk of these tumour types seen in nulliparous women compared to women with one child. By contrast, neither measure of parity was significantly associated with the risk of high-grade serous tumours (the most common histotype).

Results from previous studies and pooled analyses are broadly consistent with the greater reductions in risk of endometrioid and clear cell tumours associated with child-bearing,^{26,27,112,122} though individual studies have generally had too few cases of these fairly uncommon histotypes to show significant heterogeneity.^{108,111,119-121,213}

These analyses found no significant trend in the association between parity and

CHAPTER 7. REPRODUCTIVE FACTORS AND RISK OF OVARIAN CANCER HISTOTYPES

ovarian cancer by the age at first or last birth, which is in keeping with most studies,^{110,115,121,122,214–216} though not all.^{107,108,112,116,120,217–220}

There was some evidence of a reduced risk of ovarian cancer associated with longer durations of breastfeeding, and no significant heterogeneity of this association by histological type, consistent with some, but not all, previous studies.¹¹²

7.5.1 Mechanisms

The substantially greater reduction in risk of endometrioid and clear cell tumours seen with parity, and particularly with the first birth compared to nulliparous women, might be at least partly related to infertility. Certain conditions may both reduce a woman's fertility and increase her risk of specific histotypes of ovarian cancer. Endometriosis may be an example of one such condition, as it is an accepted cause of reduced fertility,²²¹ and pooled analyses of both retrospective¹²⁸ and prospective²⁷ studies have found that women with self-reported endometriosis had a significantly increased risk of clear cell and endometrioid cancers.

There is still considerable uncertainty regarding the mechanism for the reduced risk of ovarian cancer seen with parity and breastfeeding, assuming it is causal – other than possible confounding by causes of infertility or subfertility. Popular hypotheses include the possibilities that pregnancy and lactation act by interrupting the pro-inflammatory environment of incessant ovulation,^{140,141} by modifying the hormonal environment,²²² or by clearing pre-malignant cells from the ovary.¹⁰⁷ These analyses do not provide support either for or against these hypotheses.

7.5.2 Strengths and weaknesses

This study is one of the largest to date, allowing for more reliable comparisons of reproductive risk factors for ovarian cancer by histological type than other studies, although power was limited even in this analysis. In particular, analyses exploring the association between the timing of births and risk of ovarian cancer histotypes were under-powered, especially for the less-common histotypes.

Other strengths of the study include its prospective design, which helped minimize differential recall of exposure history amongst women with and without cancer (a potential weakness of studies with retrospective assessment of exposures), and the virtually complete follow-up and ascertainment of incident ovarian cancer, with only 1% being lost to follow-up over 14 years. One strength of the analysis approach was that the piecewise regression model allowed for a different magnitude of association for the first versus subsequent births (which more accurately reflected the data), rather than assuming a uniform relative risk per birth. Another strength of the analysis was in ensuring that analyses of breastfeeding were restricted to parous women and stratified by parity, in an attempt to remove confounding by parity as much as possible.

One weakness is the relatively short durations of breastfeeding seen in our population (of those women who reported breastfeeding, the median average duration was 2 months per child), and the lack of details regarding whether breastfeeding was exclusive to other forms of intake for the child.

7.5.3 Conclusion

In this largest prospective study of ovarian cancer to date, it has been shown that the known reduced risk of ovarian cancer with increasing parity varies between different ovarian cancer histotypes, with the strongest associations found for clear cell and en-

dometrioid tumours. The substantial increase in risk of these histotypes in nulliparous women could, at least in part, be associated with conditions that increase the risk of infertility. As any potential misclassification of tumour histotypes would tend to blur differences, the observed variation in risk by ovarian cancer histotype is unlikely to be due to confounding, and is consistent with hypotheses that the different histotypes have different aetiologies.

CHAPTER 8

IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

8.1 INTRODUCTION

Given the new hypotheses regarding the origins of serous ovarian cancer (in particular, that many high-grade serous carcinomas, the most common type of ovarian cancer, may originate from precursor lesions in the fallopian tubal epithelium²¹), there is considerable interest in investigating changes within the fallopian tubes of women with ovarian cancer.

Much of the research into the putative tubal origins of ovarian cancer has focussed on the so-called ‘serous tubal intra-epithelial carcinoma’ (STIC), and other related dysplastic processes within the tubal epithelium, which are found particularly within the fimbriae (the portion of the fallopian tube adjacent to the ovary), and from which invasive high-grade serous carcinoma is thought to arise.⁵⁰

However, as evidence for the role of STICs in the pathogenesis of high-grade serous ovarian cancer becomes more established, investigators have also begun to explore other associated changes in the fallopian tubal epithelium - both to provide further support-

ive evidence for the hypothesis of tubal origins, and because greater knowledge of pre-cancerous changes in the epithelium might aid earlier diagnosis and improved prognosis. For example, Bartlett et al. have shown evidence that women with *BRCA* mutations (who are at greatly increased risk of breast and ovarian cancer) show epigenetic reprogramming of the fimbrial epithelium, which they argue is an early pre-malignant event.²²³

8.1.1 Project investigating SOX2 expression in the fallopian tubes of ovarian cancer patients

During my DPhil, I have contributed to a project investigating the role of *SOX2*, a key stem cell differentiation gene, as a possible marker of premalignant change in the fallopian tube epithelium. Using immunohistochemistry, we have shown that *SOX2* is over-expressed in the fallopian tube epithelium of women with high-grade serous ovarian cancer, and women with germline mutations of *BRCA1/2*.²²⁴ The published paper is included in Appendix D.3.

In order to assess immunohistochemical staining within the epithelial cells in a standardised way, we wished to use a commercially-available image analysis algorithm ('Nuclear v9', from Aperio), which scores the proportion of nuclei showing immunohistochemical staining (intensity scored 1+, 2+, or 3+). However, identifying the epithelial layer (i.e. excluding nuclei from other tissue, such as stroma) requires extensive annotation of each image by hand. Given the extreme convolution of the fallopian tubal architecture, this can take up to one hour per slide. If one intends to analyse hundreds of slides, this is clearly inefficient. I was thus keen to explore potential automation of the epithelial-identification process, so as to enable future work to proceed more swiftly.

8.1.2 Existing methods of assessing immunohistochemical staining in the fallopian tube epithelium

From the literature, it appears that other researchers who have wanted to assess immunohistochemical staining of fallopian tube epithelial cells have tended either to use the same approach that we have (i.e. hand-annotation of images to identify the epithelium, followed by image analysis software to count the number and percentage of stained cells²²⁵), or have used individual researchers to assess the percentage of staining (generally taking the average of values from two independent observers²²⁶).

One group of researchers, who were interested in measuring staining in the fallopian tube epithelium of women with breast cancer, mention in their methods section that they developed an automated method of identifying the epithelium from the surrounding stroma - but no details are given of this in the paper.²²⁷ From contacting the authors, it appears that they used the commercial ‘Genie’ machine-learning pattern-recognition algorithm produced by Aperio (<http://www.leicabiosystems.com/digital-pathology/image-analysis-solutions/pattern-recognition/details/product/genie/>) (personal communication from Dr Sarah Schott).

However, the disadvantage of such an approach (quite aside from financial considerations) is that it must be trained by images that have been annotated by hand by investigators - and thus will potentially give different results in the hands of different investigators at different times. In addition, this training process could be rather time-consuming: a validation study of several image analysis approaches estimated that it took one hour of a pathologist’s time, and ten hours of a technician’s time, to train the ‘Genie’ algorithm to identify tumour areas on a single TMA of 54 cases.²²⁸ I have been unable to find an open-source (i.e. non-commercial) algorithm for identifying the fallopian tube epithelium, or one based on a standard procedure that would allow for

easy replication (i.e. not requiring machine-learning and training of the algorithm).

The literature therefore suggests that an automated method of identifying the fallopian tube epithelium would be useful to other researchers as well as our own group, and that such a method is not yet easily available or widely used.

8.2 AIM

The aim of this study is to develop and validate a more automated method than is currently available to identify the fallopian tube epithelium within scanned slide images. This will facilitate future image analysis studies of immunohistochemical staining of specific markers within the fallopian tube epithelium.

8.3 METHODS

This validation study made use of slides of fallopian tubes that had been selected, cut, stained for SOX2, and scanned, as part of a larger body of work looking at SOX2 in ovarian cancer.²²⁴ The methods below outline the creation of these scanned slide images, describe the development of a ‘macro’ (algorithm) intended to facilitate the analysis of such images, and detail the validation studies conducted on the new macro.

8.3.1 Tissue samples and slides

Anonymised pathology reports were reviewed, and unstained sections of the fallopian tubes obtained, for women treated in Oxford for high-grade serous ovarian cancer (cases), as well as women who had their fallopian tubes removed for other reasons, to act as control/ comparison groups: (a) women with benign gynaecological conditions (e.g. uterine

CHAPTER 8. IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

fibroids), (b) women with endometrial (uterine) cancer, and (c) women undergoing risk-reducing salpingoophorectomy for known inherited mutations in *BRCA1/2*.

Tissue sections were cut from formalin-fixed, paraffin-embedded (FFPE) tissue blocks at 2.5 μm or 4 μm thickness. Immunohistochemical staining was performed with the Leica Bond Max autostainer (Leica Microsystems). Slides were heated to 100°C for 20 minutes for antigen retrieval, then incubated with the rabbit anti-SOX2 (D6D9, #3579, Cell Signaling Technology), which was detected using the BOND™ Polymer Refine Detection System (DS9800, Leica Biosystems) in accordance with the manufacturer's instructions.

After staining, slides were scanned at 20x and 40x magnification on the Aperio slide scanner. For the 'gold-standard' approach for assessing nuclear staining for SOX2 within the fallopian tubal epithelium, the tubal epithelium was identified using the 'negative pen tool' in Aperio to exclude other tissues (e.g. stroma), and the staining in the marked epithelium was analysed with the commercial 'nuclear v9' algorithm in the Aperio ImageScope software (v11.2.0.780, Aperio). This algorithm assessed the intensity of nuclear staining by counting the number (and percentage) of cells that had no signal (0), weak staining (1+), moderate staining (2+) or strong staining (3+).

I contributed to the above work by reviewing the pathology reports, requesting appropriate sections, scanning stained slides and performing image analysis. Sections were cut by biomedical scientists in the Oxford Centre for Histopathology Research at the John Radcliffe Hospital, Oxford. Immunohistochemistry was performed by Leticia Campo of GCP Laboratories (part of the University of Oxford).

8.3.2 Development of a macro to identify the fallopian tube epithelium

Through a collaboration with the Computational Biology Research Group at the Weatherall Institute of Molecular Medicine (University of Oxford), a bespoke macro (algorithm) was written by Dr Dominic Waithe, for the open-source image-analysis software FIJI/ImageJ,²²⁹ to identify the epithelial layer within the fallopian tube. The image of the isolated epithelial layer can then be run through the normal nuclear-scoring algorithm.

The macro extracted the image of the epithelium in the following manner (described by Dr Waithe):

1. A low resolution version of the image is opened in FIJI/ImageJ, using the ‘third tile series’ provided by the Aperio software (corresponding to a resolution of about 2,500 by 2,000 pixels).
2. The RGB image is converted into a greyscale luminance representation (which converts Red, Green, and Blue colours into a single intensity scale).
3. The image is blurred with a Gaussian filter set to sigma=0.50, to reduce the impact of noise and small objects.
4. A Threshold is applied to the intensity image using the Moments algorithm to segment the image.
5. Thresholding produces a binary image which is then ‘despeckled’ using a median filter set to radius=0.2.
6. Small objects are then filtered from the image (area < 100).
7. The resulting mask is then applied to the original image at full resolution.

8.3.3 Validation of the macro

Validation studies were conducted to compare the macro to the current ‘gold-standard’ practice of annotating the images by hand in the Aperio image analysis software. The validation studies were in two phases: phase I compared the macro to the ‘gold standard’ of identifying the epithelium by hand, in 40 sections of fallopian tube; phase II compared the macro to both the original ‘gold standard’ of identifying the epithelium by hand, and to ‘ground-truth’ manual cell counts, in selected areas of 17 sections of fallopian tube.

8.3.3.1 Phase I

I identified 40 sections of fallopian tube (33 from patients with high-grade serous ovarian cancer, and 7 from patients who had had surgery for benign disease, such as uterine fibroids), which had been stained for SOX2. For each case, I identified the epithelial layer in two ways:

1. With image annotation by hand in the Aperio Imagescope software (using the ‘negative pen-tool’)
2. With the new macro written by Dr Waithe, run in FIJI/ImageJ

For both methods, the nuclear staining for SOX2 was assessed using the standard nuclear scoring algorithm in the Aperio Imagescope software (Aperio ‘Nuclear v9’). I then compared the assessment of staining between these methods (as indicated by the percentage of nuclei with 3+ staining), using simple descriptive statistics and Bland-Altman plots.

The time taken for the analysis process for each slide was also noted for analyses using the macro. Detailed timings were not available for analyses using the ‘gold-standard’ method, as these analyses had largely been performed prior to the idea of developing

and validating the macro. However, the approximate timing per slide had been noted informally, as it was the observation that this process was time-consuming that prompted the development of the macro.

8.3.3.2 Phase II

17 sections of fallopian tube were identified, from a mixture of patients with high-grade serous ovarian cancer and benign gynaecological conditions, which had been stained for SOX2. All cases had previously been used in the earlier validation study (phase I). For each case, the section of fallopian tube that had been used in the previous validation study was identified, and then a small area of the tube was selected - with the size of the selected area chosen so as to approximately fill an average computer screen when the image is at maximum magnification (scanned in at x20).

For each of these small selected areas, the percentage of epithelial cells that stained strongly for SOX2 with immunohistochemistry (designated '3+' in Aperio) was estimated in three ways:

1. With the epithelium identified using image annotation by hand (using the 'negative pen tool' in Aperio), and the counting of stained cells performed by the 'nuclear V9' algorithm in Aperio - the original 'gold-standard'
2. With the epithelium identified using the macro, running in FIJI/ImageJ, and the counting of stained cells performed by the 'nuclear V9' algorithm in Aperio
3. With manual counting of both stained and unstained epithelial cells in FIJI/ImageJ, using the 'multi-point' tool. This was performed separately by two observers (DW and KG).

The assessment of staining between these methods (as indicated by the percentage of nuclei with 3+ staining), was then compared using simple descriptive statistics and

CHAPTER 8. IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

Bland-Altman plots.

8.4 RESULTS

The action of the algorithm may be illustrated with one example of a section of fallopian tube: Figure 8.1 shows the normal image of the fallopian tube (with brown staining for SOX2, using immunohistochemistry), while Figure 8.2 shows the isolated epithelial layer, following use of the macro.

8.4.1 Phase I

8.4.1.1 Agreement

On average, both the macro and the hand-annotation gave fairly similar results for the percentage of tubal epithelial cells showing 3+ staining, although the percentage was slightly lower when using the macro (a mean of 11.0% (SD: 11.1%) when the epithelium was identified by the macro, compared to 13.6% (SD: 12.1%) when it was identified by hand-annotation) (Table 8.1). The ‘gold-standard’ approach estimated the 3+ staining on average to be 2.6% higher than when the epithelium was identified with the macro (95% CI: 1.5-3.7%; paired t-test: $p < 0.0001$).

This tendency for the estimate of the percentage staining to be slightly lower when

SOX2 staining (% 3+)	Method used to identify epithelium	
	By hand	Macro
Mean (SD)	13.58 (12.06)	11.01 (11.11)
Median (IQR)	12.54 (3.72-19.30)	10.84 (2.86-15.78)
Notes: SD: standard deviation. IQR: inter-quartile range.		

Table 8.1: Table comparing statistics for % of nuclei showing 3+ staining in Phase I validation sample set ($n=40$), when the tubal epithelium is identified by hand or by macro.

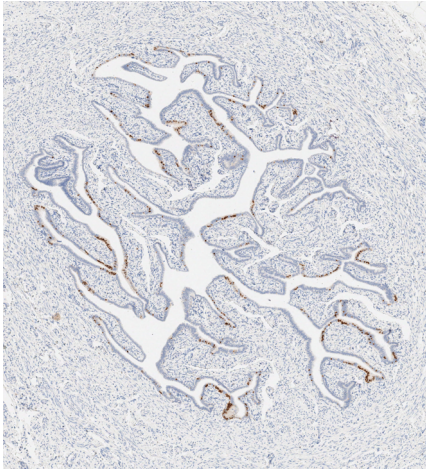


Figure 8.1: Fallopian tube section stained for SOX2

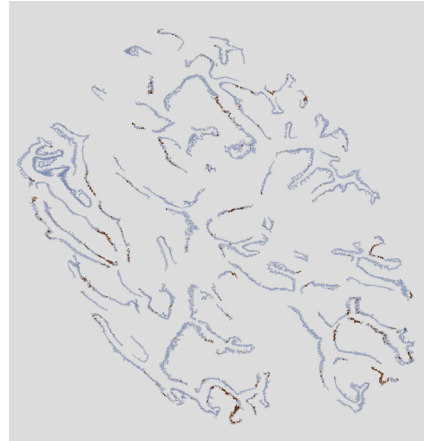


Figure 8.2: Fallopian tube epithelium, as identified by the macro

the epithelium was identified by the macro/ algorithm rather than by hand is also illustrated in a simple scatter-plot (see Figure 8.3).

However, a comparison of means (e.g. using a t-test) or a measure of the correlation is not the most appropriate measure of difference in this context. A better approach would be to compare the difference in staining estimates for each case, whilst taking account of the considerable variation in absolute values of percentage staining. This can be illustrated with a Bland-Altman plot, in which the difference between the two estimates for each case is plotted on the y-axis, against the average of the two estimates for each case on the x-axis.²³⁰

A Bland-Altman plot comparing the macro to identification of the epithelium by hand (Figure 8.4) shows that 2/40 (5%) samples were outside the limits of agreement, with one showing higher scoring, and one showing lower scoring, when the epithelium was identified by hand versus by macro. The mean difference between the two estimates was $1.6 + (0.08 \times \text{average})$ (95% limits of agreement: $\pm 2.46 \times (0.91 + 0.11 \times \text{Average})$) (units: % 3+ staining). In other words, the estimated % 3+ staining was on average about 1.6% higher when the epithelium was identified by hand (with the true value of

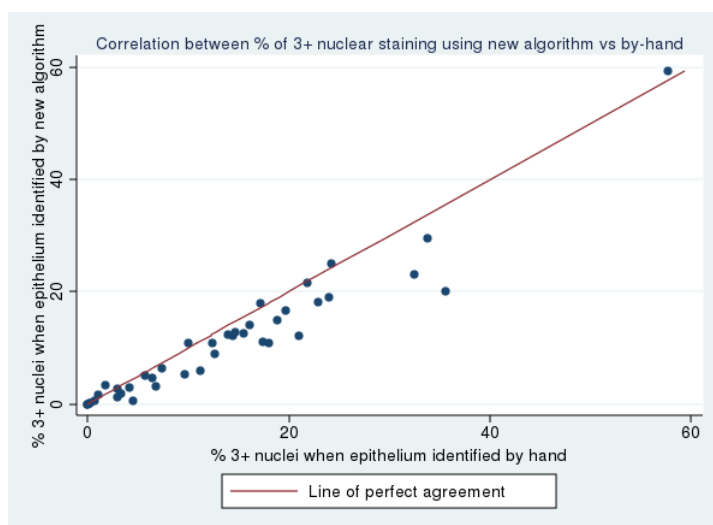


Figure 8.3: Phase I ($n=40$): Scatter plot, showing the correlation between the percentage of epithelial cells scoring 3+ for staining for SOX2, when the epithelium was identified by hand-annotation versus by the macro

this difference probably lying between about -0.9% and 4.0%), although this difference was larger for cases where the absolute value of the average % staining was higher.

The extent to which this systematic difference in estimation of staining is acceptable may depend partly on what differences in staining one wants to detect: if one is interested in substantial differences in staining between groups, then a possible under-estimation of staining by a few percent may be tolerable.

8.4.1.2 Timing

Identification of the epithelial layer was appreciably faster when using the macro than by hand-annotation. When using the macro, the whole process (including identification of the epithelium using the macro, then running the nuclear-counting algorithm) took an estimated average of 11 minutes per case (SD: 3.5). By contrast, the original ‘gold-standard’ process (using hand-annotation to identify the epithelium, followed by the nuclear-counting algorithm) tended to take about 45-60 minutes per case, and sometimes

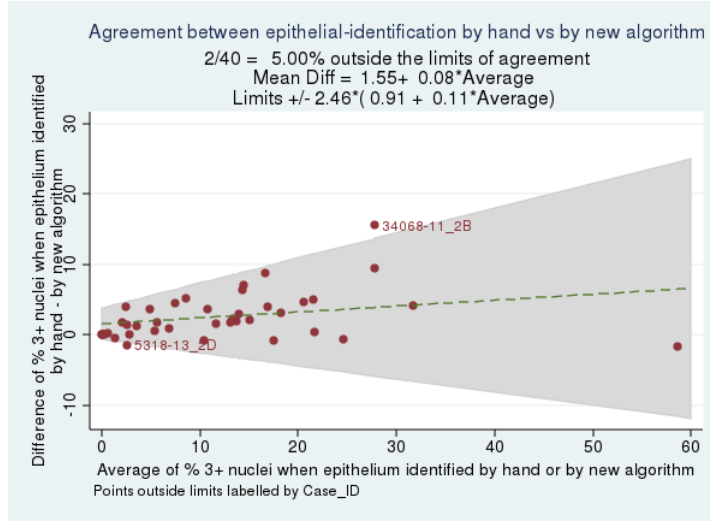


Figure 8.4: Phase I ($n=40$): Bland-Altman plot, showing the agreement between the percentage of epithelial cells scoring 3+ for staining for SOX2, when the epithelium was identified by hand-annotation versus by the macro

longer. [The estimated time taken for the ‘gold-standard’ method is more approximate, as precise timings were not recorded for these initial annotations].

8.4.2 Phase II

8.4.2.1 Agreement

On average, all four methods (hand-annotation in Aperio, use of the macro, or hand-counting of cells by two observers) gave fairly similar results for the percentage of tubal epithelial cells showing 3+ staining. While there were some differences, these tended not to be statistically significant - partly due to the small number of cases in the study, and the wide distribution (see Table 8.2).

A simple scatter plot to compare the agreement between the two observers (KG and DW) in the *total number of cells* counted is seen in Figure 8.5. The plot shows that, while the two attempts to count cells manually produced highly correlated results, one

CHAPTER 8. IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

	Manual (KG)	Method used to identify epithelium Manual (DW)	Macro	Annotation
SOX2 staining (% 3+)	Manual (KG)	Method used to count stained cells Manual (DW)	Aperio	Aperio
Mean (SD)	23.01 (15.96)	21.34 (14.43)	19.35 (15.56)	21.51 (15.93)
Median (IQR)	20.79 (11.24-34.18)	16.42 (12.81-28.57)	18.31 (5.55-25.33)	25.43 (8.16-31.99)

Notes: SD: standard deviation. IQR: inter-quartile range.

Table 8.2: Table comparing statistics for % of nuclei showing 3+ staining in Phase II validation sample set ($n=17$), when the stained tubal epithelial cells are identified and counted either manually by individuals (KG & DW), or with tubal epithelium identified by a macro or hand-annotation in Aperio, and the counting of stained cells performed by the ‘nuclear V9’ algorithm in Aperio

observer (KG) tended to count fewer cells than the other observer (DW).

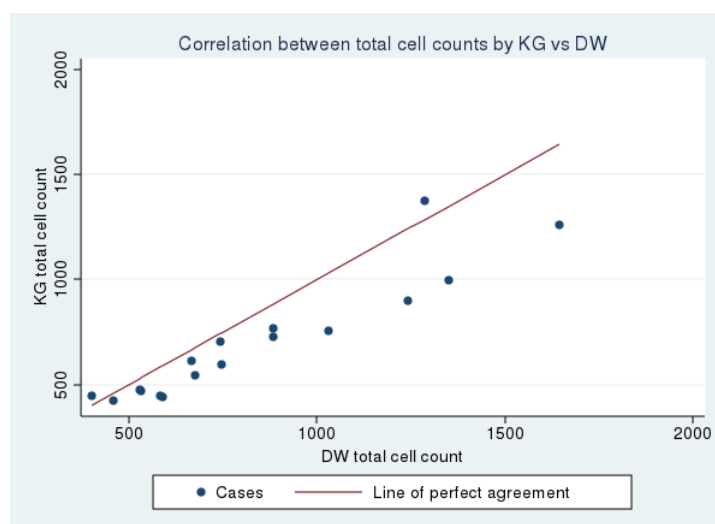


Figure 8.5: Phase II ($n=17$): Scatter plot showing the correlation between two observers (KG and DW) in counting total number of epithelial cells

However, a scatter plot comparing the agreement between the two observers in the *percentage of strongly-stained cells* (seen in Figure 8.6) shows that there was fairly good agreement between the observers in the percentage of cells stained, despite the systematic differences in counting absolute numbers of cells.

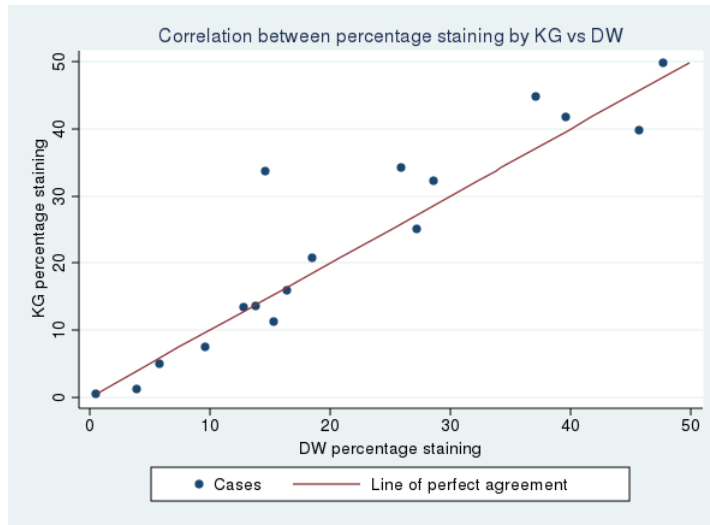


Figure 8.6: Phase II: Scatter plot showing the correlation between two observers (KG and DW) in counting the percentage of epithelial cells stained for SOX2 by immunohistochemistry

For subsequent analyses, an average was taken of the manual counts of percentage staining by the two observers for each case. This average was then compared to the two other methods (when the epithelium was identified by hand-annotation in Aperio, vs identified by the FIJI/ImageJ macro).

Bland-Altman plots were used to compare the average percentage staining from manual cell counts (by KG & DW) to the percentage ‘3+’ staining as judged by the Aperio algorithm, both when the epithelium was identified by hand-annotation in Aperio (Figure 8.7), and when the epithelium was identified by the macro (Figure 8.8). A Bland-Altman plot was also used to compare the agreement in average percentage ‘3+’ staining when the epithelium was identified by hand-annotation in Aperio, vs by the macro, with counting done by the Aperio algorithm in both cases (Figure 8.9).

8.4.2.2 Investigating outliers

Although there tended to be fairly good agreement, there were some cases with substantially higher percentages stained when one measure of staining was used compared to

CHAPTER 8. IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

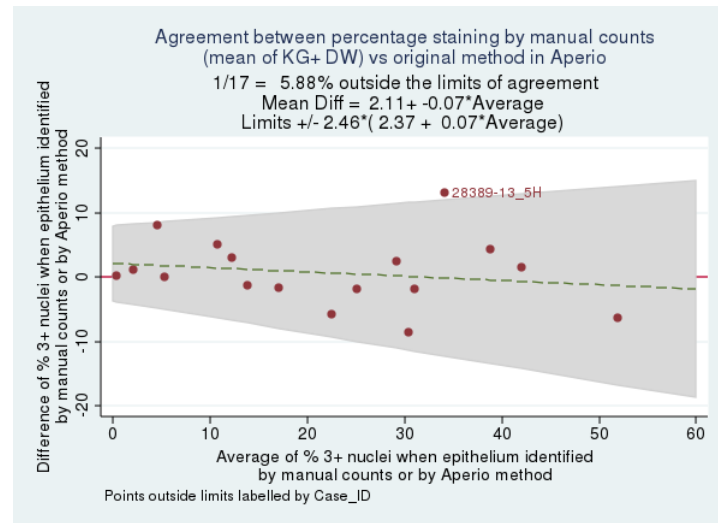


Figure 8.7: Phase II: Bland-Altman plot, showing the agreement between the percentage of epithelial cells scoring 3+ for staining for SOX2, comparing a) manual counts vs b) the original ‘gold-standard’ method of identifying the epithelium by hand-annotation in Apero, followed by automated cell-counting in Apero

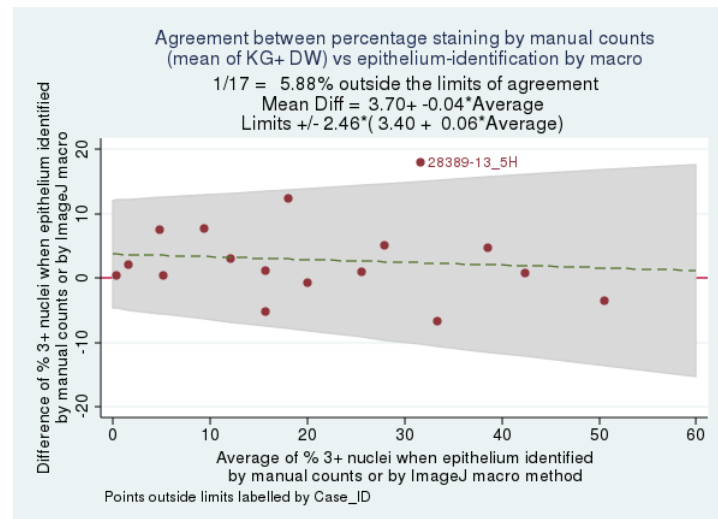


Figure 8.8: Phase II: Bland-Altman plot, showing the agreement between the percentage of epithelial cells scoring 3+ for staining for SOX2, comparing a) manual counts vs b) epithelial identification by the macro, followed by automated cell-counting in Apero

CHAPTER 8. IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

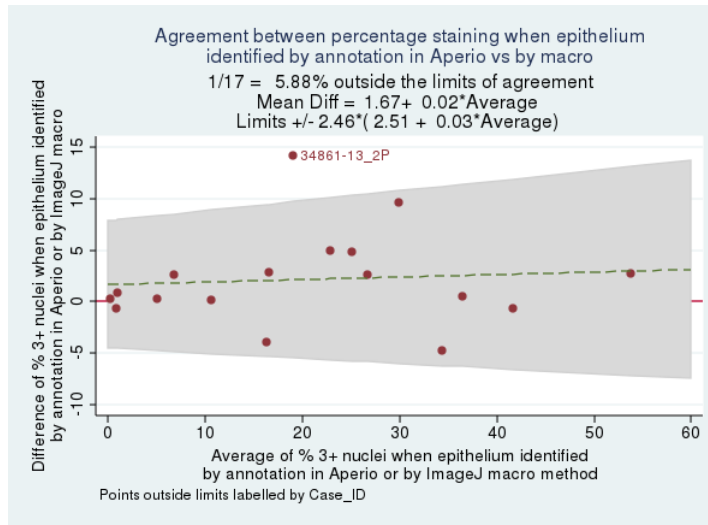


Figure 8.9: Phase II: Bland-Altman plot, showing the agreement between the percentage of epithelial cells scoring 3+ for staining for SOX2, comparing a) identifying the epithelium by hand-annotation in Aperio, followed by automated cell-counting in Aperio vs b) epithelial identification by the macro, followed by automated cell-counting in Aperio

another. Two of these are investigated below.

Case 34861-13-2P

Case 34861-13_2P was an outlier in Figure 8.9, with a much higher percentage of staining recorded when the epithelium was identified by hand-annotation in Aperio than when the epithelium was identified by the macro.

On examining the images, it is possible to see why this discrepancy arose. Figure 8.10 shows the image with the epithelium identified by hand-annotation, while Figure 8.11 shows the epithelium identified by the macro. The macro appears to have missed one area of epithelium on the right-hand side of the image, which included many stained cells. In addition, there was a group of detached, unstained epithelial cells in the lumen of the tube, which were edited out by hand-annotation, but understandably persisted in the macro image. The combination of these two factors (which reduced the number of stained cells, and increased the number of unstained cells, in the macro image relative to the hand-annotation image) explains the discrepancy observed.

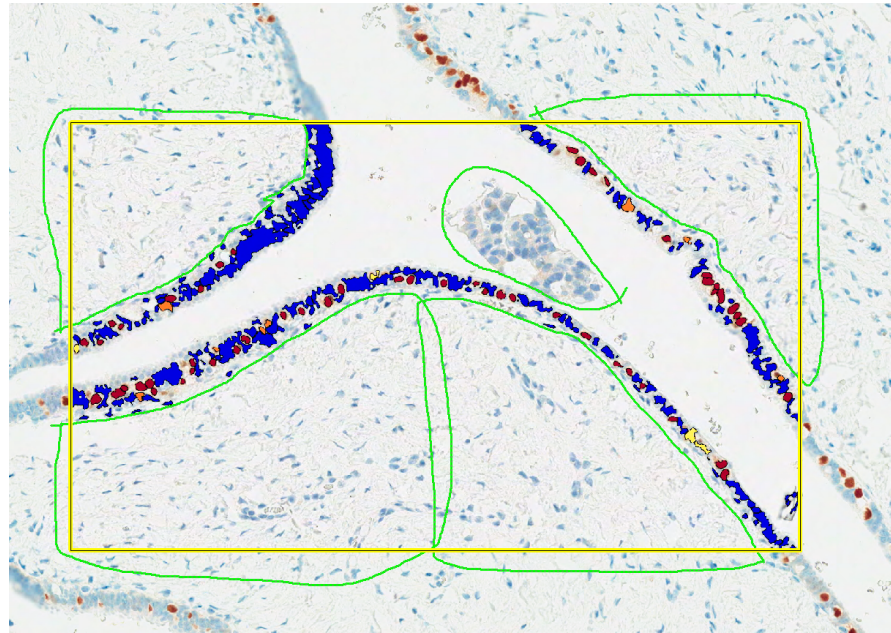


Figure 8.10: Fallopian tube section stained for SOX2, with the epithelium identified by hand-annotation, and staining assessed by Aperio software. This is a false-colour image, illustrating staining as assessed by Aperio: unstained nuclei are marked in blue, while nuclei stained with SOX2 are marked in yellow (1+), orange (2+), or brown (3+).

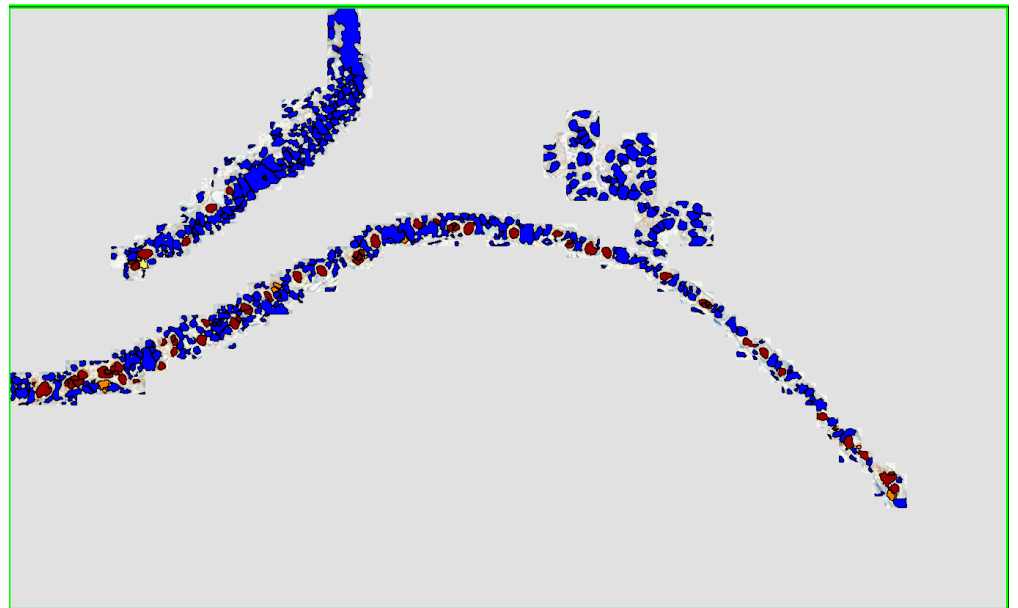


Figure 8.11: Fallopian tube section stained for SOX2, with epithelium identified by the macro in FIJI/ ImageJ, and staining assessed by Aperio software

Case 28389-13-5H

Case 28389-13_5H was an outlier in both Figure 8.7 and Figure 8.8, in which the mean percentage of stained cells as judged by manual counts by the two observers was about 15% higher than that measured by the Aperio algorithm (regardless of whether the epithelium had been identified by annotation or by the macro).

On examining the images in Figures 8.12, 8.13, 8.14, and 8.15, it is not immediately clear why this discrepancy arose. Presumably the observers performing manual counts either over-counted stained nuclei, or under-counted unstained nuclei, relative to the Aperio software.

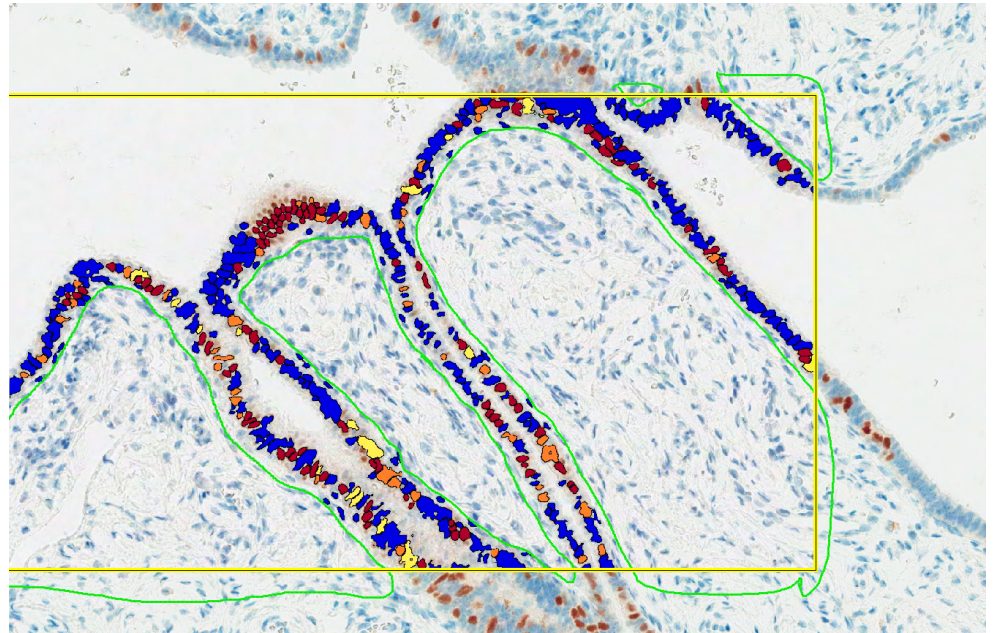


Figure 8.12: Fallopian tube section, stained for SOX2, with epithelium identified by hand-annotation, and staining assessed by Aperio software

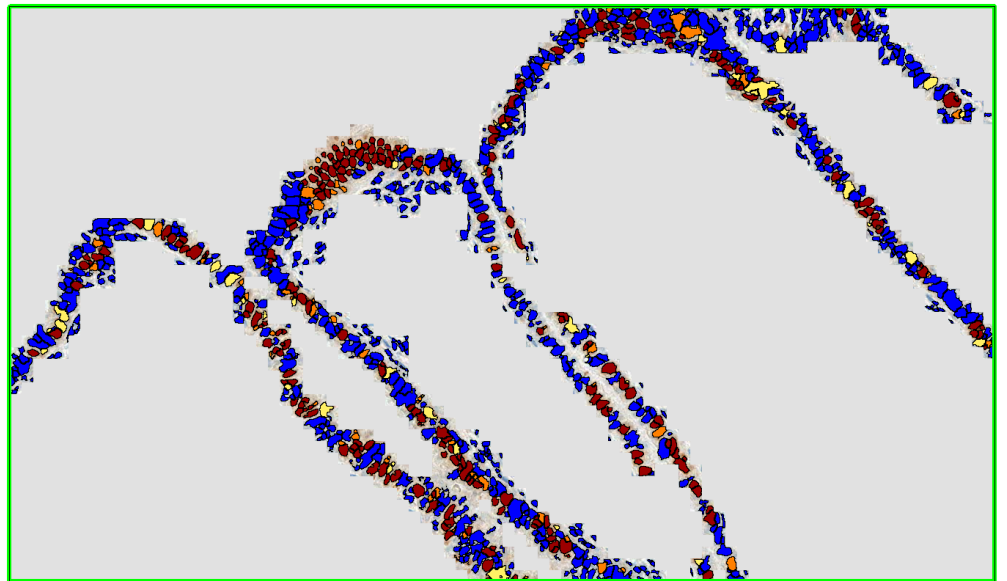


Figure 8.13: Fallopian tube section, stained for SOX2, with epithelium identified by macro in FIJI/ImageJ, and staining assessed by Aperio software

CHAPTER 8. IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

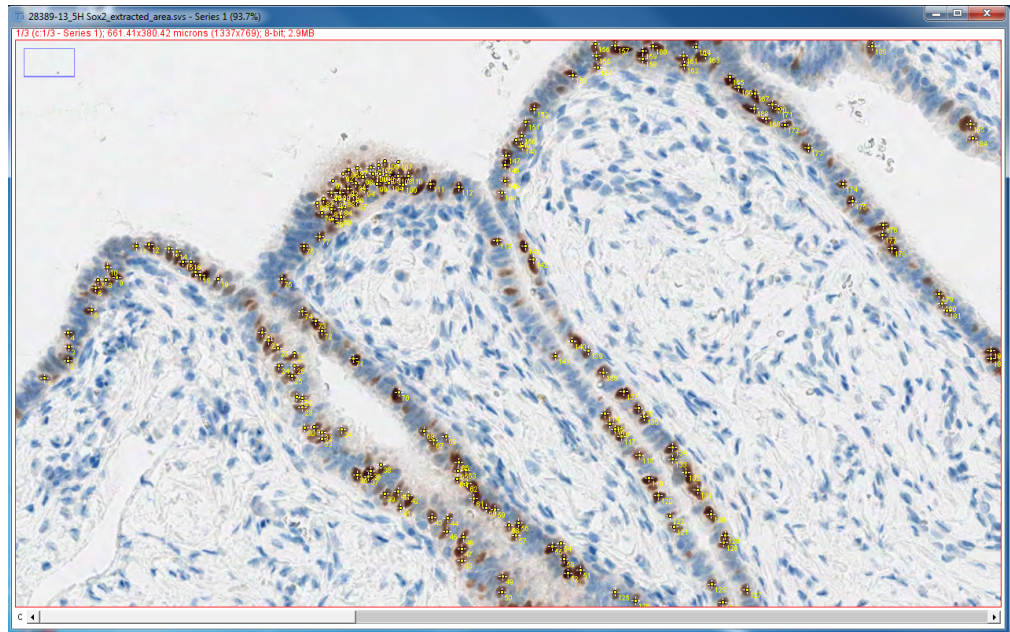


Figure 8.14: Fallopian tube section, stained for SOX2, with stained epithelial nuclei marked by manual counts in FIJI/ImageJ

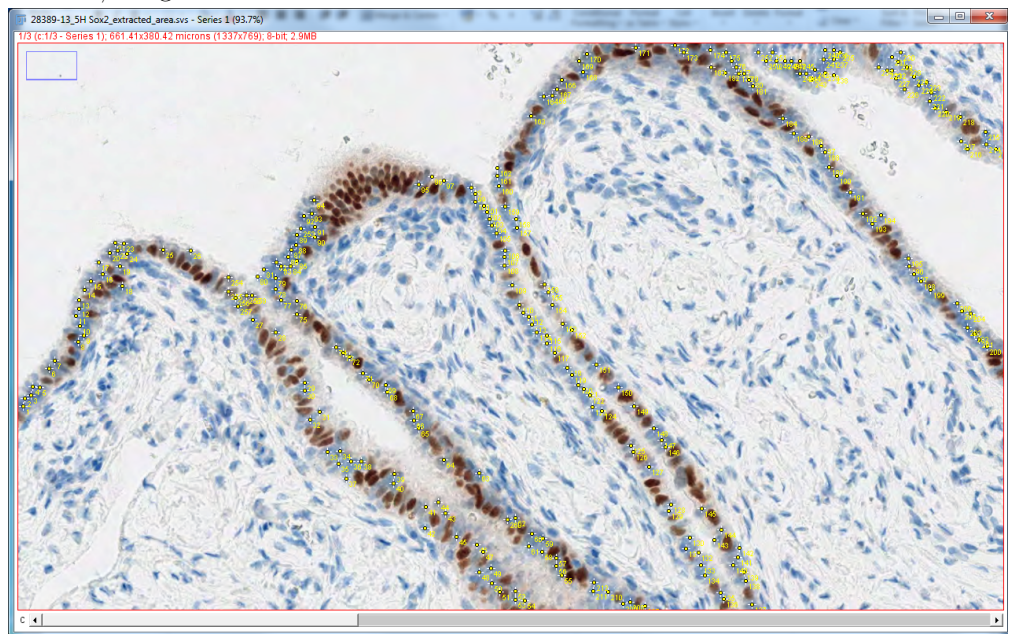


Figure 8.15: Fallopian tube section, stained for SOX2, with unstained epithelial nuclei marked by manual counts in FIJI/ImageJ

CHAPTER 8. IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

SOX2 staining (% 3+)	Absolute difference in percentage staining	
	Between two observers performing manual cell-counts	Between Aperio-annotation and macro method
Mean (SD)	3.79 (4.67)	3.33 (3.75)
Median (IQR)	2.26 (0.78-4.06)	2.63 (0.6-4.72)
Notes: SD: standard deviation. IQR: inter-quartile range.		

Table 8.3: Table comparing statistics for absolute differences in the % of nuclei showing 3+ staining in validation sample set ($n=17$), comparing differences between two human observers, and the differences when the tubal epithelium is identified by a macro vs hand-annotation in Aperio

8.4.3 Comparing inter-observer vs inter-method variation

One approach to evaluating whether it is reasonable to use the macro is to consider whether any differences seen between the measure of percentage staining when the epithelium is identified with the macro, compared to when it is identified with hand-annotation in Aperio (the ‘gold-standard’), are greater than any differences between human observers performing manual cell-counts.

Table 8.3 shows some summary statistics, comparing two sets of absolute differences in the percentage of staining:

- The absolute difference in percentage staining, when the stained and unstained cells were counted manually by two observers (i.e. inter-observer differences)
- The absolute difference in percentage staining when the counting of stained cells was performed by the Aperio ‘Nuclear V9’ algorithm, comparing when the epithelium was identified by hand-annotation in Aperio, vs when it was identified by the FIJI/ImageJ macro.

As can be seen, the differences appear quite similar. Formal statistical testing (with the non-parametric Wilcoxon matched-pairs signed-ranks test) suggests that the

CHAPTER 8. IMAGE ANALYSIS OF THE FALLOPIAN TUBES: MACRO DEVELOPMENT AND VALIDATION

inter-observer differences between two humans performing manual cell counts are not significantly different to the differences arising between the two methods of identifying the epithelium ($p=0.6$).

8.5 DISCUSSION

Taken together, this two-phase validation study shows that there is reasonable agreement between the scoring of nuclear staining when the epithelium is identified by the macro compared to annotation by hand. Interestingly, there appears to be a slight but consistent under-scoring when the epithelium has been identified by the macro. However, this difference appears to be smaller than the differences between two human observers performing manual cell counts.

Importantly, the analysis process took an average of only 11 minutes per slide with the macro, compared to an estimated 45-60 minutes with the original hand-annotation method. Thus, although there is not perfect agreement between assessment of staining when the epithelium is identified by the macro versus by hand-annotation, it may be worth using the macro in subsequent image-analysis work, due to the potential for significant time-saving.

Indeed, the macro is already being put to use in research by Fiona Chen (a visiting medical student, whose research project I am helping to co-supervise with Prof Ahmed), as she is using immunohistochemistry to investigate differential staining for specific markers in the fallopian tube epithelium of women with high-grade serous ovarian cancer and other ovarian tumours. As she wishes to analyse staining with three different markers in sections of fallopian tubes from about 70 patients, this volume of analysis work would not be feasible for one person to complete, using the hand-annotation method - and so the macro (with its reduced analysis time) has proved invaluable.

CHAPTER 9

MWS OVARIAN CANCER TISSUE PILOT STUDY

9.1 INTRODUCTION

The Million Women Study did not collect biological samples (such as blood or tissue) as part of the original protocol. Subsequently, the Disease Susceptibility Study was established: this is a nested case-control study, involving collection of blood samples from about 20,000 women with incident breast cancer or cardiovascular disease (cases), and about 40,000 controls (women without incident breast cancer or cardiovascular disease), to look at potential biomarkers or genetic markers associated with disease risk. However, tumour tissue samples have not been collected.

Given the substantial heterogeneity between the different histological types of ovarian cancer, not only in terms of morphological appearance, but also in terms of molecular characteristics and putative origins (see Chapter 2), and risk factors (see Chapters 3 and 5 to 7), it would obviously be advantageous to obtain samples of ovarian cancer tumour tissue, to facilitate further analyses. In particular, samples of tumour tissue could be used to create tissue micro-arrays for immunohistochemistry, for more detailed molecular characterisation of the tumours.

Therefore, a pilot study was planned, to explore the feasibility of identifying, obtaining, and processing tissue samples from participants in the Million Women Study who have been diagnosed with ovarian cancer. The pilot study is initially restricted to women with ovarian cancer diagnosed in Oxford, partly for simplicity of logistics (as there are some shared governance procedures in place between the University of Oxford and the Oxford University NHS Hospitals Foundation Trust), but also because the Oxford hospitals represent one of the largest single centres in terms of the number of MWS participants treated for ovarian cancer.

9.2 AIMS

The initial aims of the pilot study were:

- To apply for and obtain ethical approval for the pilot study
- To verify that it is possible to correctly identify participants in the Million Women Study who were treated for ovarian cancer in a specific centre, using hospital records
- To obtain and review the pathology report, and identify what samples are available for such women
- To develop an initial protocol for obtaining and processing relevant tissue samples.

Assuming that the above aims were successful, further aims included:

- To retrieve and process tissue samples in an appropriate manner
- To identify challenges and potential problems, and modify the protocol accordingly for future work
- To use the data collected (both electronic and tissue samples) to conduct some initial validation studies.

If all of the above aims are successful, then the hope would be to expand the pilot study to collect tissue samples of MWS participants treated for ovarian cancer in other centres. However, due to financial, timing, and logistical constraints, this is beyond the scope of this DPhil project.

9.3 METHODS

This pilot study was, of necessity, a highly collaborative endeavour, involving contributions from a range of researchers, clinicians, and administrators.

9.3.1 Initial planning and ethical approval

The initial plan for the pilot study and application for ethical approval were written with input from Professor Jane Green, Professor Dame Valerie Beral, and Professor Ahmed Ahmed, and with extensive practical assistance from Krys Baker (Research Co-ordinator for CEU) and Lynden Guiver (Study Co-ordinator for MWS). The approved amendment to the MWS protocol (Appendix 28 of the Million Women Study) is included in Appendix B.1.

9.3.2 Research approval from the Oxford Centre for Histopathology Research (OCHRe)

The application for approval of collaboration with the Oxford Centre for Histopathology Research (OCHRe) was written with input from Professor Ahmed and Professor Green, with practical assistance from Lynden Guiver.

9.3.3 Development of an initial protocol for tissue processing

The initial planned protocol for tissue processing was developed with input from Professor Ahmed, and following discussions with Professor Gillian Reeves, and Associate Professor Ruth Travis, both of whom have developed protocols for retrieving and processing archived tumour tissue (for breast cancer and prostate cancer, respectively). These protocols were adapted and simplified for the current project of investigating ovarian cancer, as the pathological and clinical differences between cancer types inevitably affect the design of the protocols. For example, ovarian cancer does not have large areas of in-situ disease that would require separate sampling to the invasive disease (cf. breast cancer).

However, the hypothesis that some ovarian cancers may originate in the fallopian tubes means that it would be advisable to obtain samples of the fallopian tubes. I initially considered attempting to take cores of fallopian tube tissue, to generate a tissue micro-array of the fallopian tubes. However, given that the putative precursor lesions for high-grade serous ovarian cancer (e.g. serous tubal intraepithelial carcinoma, STIC) tend to be small and focal when identified, I decided that random sampling of the fallopian tube epithelium would probably be unhelpful. More targeted sampling was considered, but there were concerns that this might deplete a valuable resource. It was therefore decided to merely take unstained sections of the tubes, which would allow limited immunohistochemistry.

9.3.4 Identification of participants and data transfer and linkage

MWS participants with ovarian cancer (coded C56 in ICD-10) recorded for a hospital admission in the Oxford region were identified using hospital records by Carol Hermon, a data analyst for the Million Women Study. A list was generated of patient identifiers (NHS number and date of birth) for these women, with a study-specific ID. This list was

securely encrypted and transferred by email to the Oxford Centre for Histopathology Research (OCHRe); the passcode was given by telephone.

OCHRe linked the list to the electronic histopathology record, and exported selected sections of the electronic pathology reports into a spreadsheet, which was stripped of patient identifiers other than the study-specific ID. This file was encrypted and sent back to the Cancer Epidemiology Unit by email; the password was given separately.

This process was done in two batches: the first batch were women with a hospital admission with a diagnosis of ovarian cancer from January 2014 onwards ($n=28$); the second batch were those with a hospital diagnosis up to December 2013 ($n=258$ for a diagnosis at an Oxford hospital, of which $n=103$ for a procedure code indicating likely suitable samples, e.g. hysterectomy, salpingoophorectomy, or salpingectomy). This batching was done for practical reasons: OCHRe currently stores blocks and slides from cases prior to January 2014 in an off-site archive, whereas those from January 2014 onwards are stored in Oxford. It therefore made sense to begin the pilot study by focussing on specimens that are likely to be easily retrievable.

9.3.5 Review of histopathology reports and identification of potential specimens

I reviewed the histopathology reports. Women were excluded if tissue specimens were clearly unavailable in Oxford (e.g. if slides had just been sent to Oxford for histopathology review, and then returned to the sender). For each woman, the likely best sample was identified.

Samples were excluded if they were not of an ovarian cancer (e.g. some women also had resections of skin cancers), or if they did not contain viable tumour (e.g. post-chemotherapy samples with complete response, and no viable residual tumour identi-

fied). Primary biopsy specimens could be considered if resection specimens (e.g. salpingoophorectomy +/- hysterectomy) were reported to not contain viable tumour.

9.4 RESULTS

Ethical approval for linkage with the pathology reports was covered by the original ethical approval for MWS, when women were consented for access to their medical record. Ethical approval for the ovarian cancer tissue pilot study was granted by the ‘East of England - Cambridge South’ Research Ethics Committee on 4th March 2016 (Million Women Study Amendment 28) (see Appendix B.2). Initial approval for linkage with OCHRe was granted in January 2016 (reference: 16/A014).

Below are the results of Phase I of the pilot study: linkage to the histopathology records.

9.4.1 Batch 1: Admissions from January 2014

28 women were identified as MWS participants who had an admission with ovarian cancer in Oxford since January 2014, according to hospital records. This list was linked with the pathology database at OCHRe, and the electronic histopathology reports obtained, as described above.

9.4.1.1 Women with incident ovarian cancer in cancer registry data

Of these 28 women, 10 (36%) were registered as having incident ovarian cancers according to cancer registry data (from ONS and NCIN/PHE) (see Table 9.1). For these 10 women, the pathology reports confirmed that 7 had a diagnosis of ovarian cancer, and 2 had cancers that were thought more likely to arise from the peritoneum or fallopian tube

(but could have been ovarian). One woman did not have a recorded diagnosis of ovarian cancer in the pathology report, but information elsewhere in her hospital record confirmed that she had been diagnosed with ovarian cancer elsewhere, and subsequently admitted to Oxford for treatment.

The histological type of ovarian cancer from the cancer registry data accorded with that from the the pathology reports in 8/9 cases. The one slight discordance was for a carcinosarcoma with a high-grade serous epithelial component, which was recorded as high-grade serous ovarian cancer in the cancer registry data.

The pathology reports suggested that 9/10 of these cases should have potentially usable pathology samples (6 resection specimens, 3 biopsy-only); the case diagnosed outside of Oxford had no suitable specimen available. For 2/3 cases with only biopsy samples available, this was because no resection specimen was recorded; for 1/3, a resection specimen was available, but there had been a complete response to pre-operative chemotherapy, and no viable tumour was seen. Thus, based on this small initial batch of pilot study data, one might expect about 90% of incident ovarian cancer cases to have potentially-suitable samples available.

9.4.1.2 Women without incident ovarian cancer in cancer registry data

Data were also explored for the 18/28 women who had a diagnosis of ovarian cancer in electronic hospital records data, but who were not incident ovarian cancers according to cancer registry data (see Table 9.1). 3/18 women did have a confirmed diagnosis of ovarian (or fallopian tube or peritoneal) cancer in the pathology report, but also had a previous incident cancer at another site. 6/18 women were listed as having primary peritoneal or fallopian tube (rather than ovarian) cancers in cancer registry data (these have not generally been considered ‘ovarian’ cancers for the purpose of this thesis, but arguably their similarities in morphology and management, the difficulty in distinguishing

them, and their putative shared origins mean that they should be considered together in the future). 5/18 women had evidence of a cancer at another site (2 lung, 2 endometrial, and one breast cancer), in both cancer registry data and pathology reports, and no mention of ovarian cancer in the pathology reports. 2/18 women had ovarian (or fallopian tube) cancers confirmed in the pathology report, but diagnosed after the end of follow-up for cancer incidence in cancer registry data (as we currently have hospital record data up to a more recent date than cancer registry data). A further 2/18 women had ovarian cancers confirmed in the pathology report but no record of an incident cancer in cancer registry data, and no clear reason for this omission - though in both cases the date of the pathology specimen being taken was close to the end of follow-up for cancer registry data, and so even a relatively short time-lag in diagnosis and registration could explain this.

Thus, overall there were 10 additional cases with apparent incident cancers of the ovaries, peritoneum or fallopian tube - which were not considered incident ovarian cancers by the usual cancer-registry definition used in this thesis, either due to site (fallopian tubal or primary peritoneal) or time of diagnosis (after/near the end of follow-up for cancer registry data), but which might be worth including in subsequent tissue-based analyses. Of these, the pathology reports suggested that 9/10 should have suitable tissue samples (8 resection specimens, 1 biopsy only); 1/10 only had records of slides that had been received for review and then returned.

CHAPTER 9. MWS OVARIAN CANCER TISSUE PILOT STUDY

Table 9.1: *Batch 1: Table showing the results of comparing women with a diagnosis of ovarian cancer in electronic hospital records to data from cancer registries and pathology reports*

Result of comparison	n	%
Confirmed incident ovarian cancer	10	36
Previous incident cancer	3	11
Site listed as peritoneal or fallopian tube in cancer registry data	6	21
Cancer at other site in both cancer registry and pathology record	5	18
Cancer diagnosed after end of cancer registry follow-up	2	7
Cancer diagnosed near end of cancer registry follow-up	2	7
Total	28	100

Notes: The above results are for batch 1 (i.e. women with a hospital admission with a diagnosis of ovarian cancer in the electronic hospital records from January 2014 to the end of follow-up for hospital records).

9.4.2 Batch 2: Admissions up to December 2013

258 women were identified as MWS participants who had a diagnosis of ovarian cancer recorded at an Oxford hospital up to December 2013, of whom 103 also had a recorded procedure code for a likely relevant resection specimen (e.g. hysterectomy, salpingoophorectomy, or salpingectomy). 17/258 were removed by standard exclusions (e.g. previous cancers before recruitment). Of those remaining, 189/241 (78%) women matched to the Oxford pathology records database (i.e. they had at least one identified pathology report). Of those women who matched to the pathology record, 139/189 (74%) had an incident diagnosis of ovarian cancer according to cancer registry data. Detailed review of the pathology reports from Batch 2, to compare diagnoses in the pathology

report to those from cancer registries and hospital records, is still ongoing. However, if we assume that the first batch of cases is representative, then we might expect that about 90% of incident ovarian cases would have potentially-usable tissue samples - i.e. $0.9 \times 139 = 125$ cases.

Formal approval of the second phase of the pilot study (i.e. actual retrieval of archived tissue blocks and slides) by OCHRe is still awaited, and thus tissue retrieval has not yet commenced.

9.4.3 Planned protocol for slide review and sample processing

Part of the aim of this pilot study was to develop an initial protocol for sample identification and processing. This protocol is thus outlined below, even though capacity constraints at OCHRe have meant that it has not yet been possible to begin retrieval of slides and samples. The protocol aims to ensure that samples are taken from both primary and metastatic tumour deposits (thus allowing for comparisons between the two), as well as unstained sections from the fallopian tubes. Given that cores will be taken for tissue micro-arrays, we also wish to consider taking cores for possible future DNA extraction for targeted sequencing - though this is obviously exploratory, and dependent on having reliable DNA extraction methods for formalin-fixed, paraffin-embedded tissue.

The protocol is flexible, and will need to take account of the amount of tissue available (i.e. ensuring that samples are not depleted, and hence not taking cores from single biopsy specimens). In addition, the protocol makes use of existing interest in digital image analysis, by ensuring that key slides are scanned and stored for possible future projects.

The original diagnostic slides, stained with haematoxylin and eosin (H & E), will be reviewed by a pathologist (either myself or a collaborating consultant pathologist).

Slides will be identified for scanning, and for sampling for tissue micro-arrays.

Slides identified for scanning will be:

- Any slides of the fallopian tubes
- Any slides from which cores are to be taken for tissue micro-arrays.
- A slide showing endometriosis, if present
- For biopsy-only specimens: the main biopsy slide

Slides identified for sampling from the tissue blocks will be:

- Cores for tissue micro-arrays (TMAs): both primary tumour and secondary/ metastasis, if available (e.g. tumour in the omentum)
- Unstained sections x10 from each fallopian tube
- Cores for possible future DNA extraction: of invasive tumour (both primary and secondary if available) and normal tissue
- For biopsy-only specimens: x5 unstained sections (if possible)

The plan would be that staff at OCHRe would then scan the selected slides, and cut sections and take tissue cores for tissue-microarrays, and for storage for possible future DNA extraction, as appropriate. Immunohistochemistry (on both the fallopian tube sections and the TMAs of tumour tissue) would be performed at a collaborating laboratory in Oxford.

9.5 DISCUSSION

The first phase of the pilot study has established procedures for identifying MWS participants with ovarian cancer who were treated in a specific centre, and for securely linking to pathology databases to obtain electronic copies of the pathology reports. Review of these

reports has allowed for validation of diagnoses received from routine data sources, as well as providing more detailed information regarding the likely tissue specimens available. This is helping to inform the second phase of the pilot study (in which tissue samples are to be retrieved, reviewed, and processed), which is still ongoing.

9.5.1 Challenges

Several important issues, both logistic and scientific, were raised by the first phase of the pilot study. These included:

1. **Incorrect information from hospital records**

Some women with a diagnosis of ovarian cancer in data from the electronic hospital records proved not to have evidence of ovarian cancer on detailed review of the pathology reports. For example, there was a case for which the primary bulk tumour had been found on the ovary, and a biopsy performed, which confirmed malignancy - so the patient was diagnosed with ovarian cancer. However, a subsequent hysterectomy showed that the primary in fact appeared to have arisen from the endometrium (uterus). This alteration in the diagnosis of the primary site occurred some time later, and thus the original information was presumably entered in the hospital record, without correction. However, the information from cancer registry data for this case was correct (endometrial cancer). Thus, it may be that a combination of case identification from cancer registries and hospital records, with confirmation of the diagnosis from pathology reports, would be most reliable.

2. **Storage of samples off-site**

In the Oxford University NHS Foundation Trust, histopathology samples (blocks and slides) are moved to an off-site storage facility after 2-3 years. For example, for a pilot study conducted in January-September 2016, samples were only available within Oxford for cases from January 2014 onwards. Retrieving samples from the

off-site facility is possible, but entails increased costs of both time and money.

3. Availability of suitable samples

Even for those cases of ovarian cancer diagnosed in Oxford, not all had resection specimens with abundant viable tumour. In some cases, only biopsies were performed; in others, pre-operative (neo-adjuvant) chemotherapy was so effective that no viable tumour was seen in the resection specimen. Thus, protocols for tissue processing will have to take account of varying tissue availability.

9.5.2 Conclusion

The first phase of this pilot study has established that it is feasible to identify MWS participants with ovarian cancer treated in a particular area or centre, and then link securely to local histopathology databases in order to obtain reports and identify samples. Review of the pathology reports allows for validation of the diagnoses recorded in data from electronic hospital records and cancer registries - both in terms of cancer site and tumour morphology - generally showing fairly good agreement. This first phase has also identified some potential challenges for the second phase of the study, including the potential costs (both time and expense) of retrieving older samples from off-site archives, and the possibility that some cases may not have samples that are accessible and usable.

CHAPTER 10

DISCUSSION

10.1 OVERVIEW

10.2 AIM OF MY DPHIL WORK

My main aim was to investigate risk factors for ovarian cancer by histological tumour type, focussing in particular on factors previously identified as being associated with ovarian cancer by other studies, and which might relate to new hypotheses of the origins of ovarian cancer, but which had not previously been reported on by either the Million Women Study or the Collaborative Group on Epidemiological Studies of Ovarian Cancer. Specifically, the factors identified for further investigation were tubal ligation (with regard to risk of both ovarian cancer and other related or common cancers) and reproductive factors (parity and breastfeeding).

A secondary aim was to develop resources for further investigation of the origins and histological types of ovarian cancer: by developing and validating methods to assist in image analysis of the fallopian tubes (of interest as a putative site of origin for the most common type of ovarian cancer), and by conducting a pilot study of obtaining pathology

reports and tissue samples from participants in the Million Women Study with ovarian cancer.

10.3 MAIN FINDINGS FROM THIS THESIS

The main findings of this thesis, as described in Chapters 5 to 9 are summarised below.

10.3.1 Tubal ligation and risk of ovarian cancer histotypes

The analyses in Chapter 5 confirmed that, overall, women who have had a tubal ligation have about a 20% reduction in risk of ovarian cancer - but that this varies significantly by histotype. The greatest reduction in risk was seen for endometrioid and clear cell tumours, for which the risk was about halved. There was about a 20% reduction in risk of high-grade serous carcinomas, which was statistically different to the null association with low-grade serous tumours (borderline and low-grade serous carcinoma). There was no association seen with mucinous tumours.

These findings are consistent with the new hypotheses of possible extra-ovarian origins for certain ovarian cancer histotypes. The observation that tubal ligation is associated with a substantial reduction of endometrioid and clear cell tumours would be consistent with the hypothesis that some of them arise from endometriosis, as tubal ligation could potentially block the passage of cells from the endometrium through the fallopian tubes (a potential mechanism for endometriosis). The reduction in risk of high-grade serous carcinomas could be consistent with their putative origins in precursor lesions in the fallopian tubes, which might be disrupted by tubal ligation surgery.

There was, however, little evidence of an effect of timing of tubal ligation - but the study was probably statistically under-powered to detect a true effect of timing of tubal ligation by histological type, if it was there.

10.3.2 Tubal ligation and risk of other cancers

The analyses in Chapter 6 expanded those in the previous chapter on tubal ligation and ovarian cancer risk, by exploring the association between tubal ligation and the risk of an additional 25 cancer types. The results showed that tubal ligation is not only associated with about a 20% reduction in the risk of ovarian cancer, but is also associated with a similar-magnitude reduction in the risk of peritoneal cancer, and a significant reduction in the risk of fallopian tube cancer. By contrast, tubal ligation was not associated with reductions in the risk of other cancer types, including hormonally-driven cancers such as those of the breast and endometrium.

These findings are of interest, because they shed further light on the possible origins of ovarian cancer, and also of related cancers. High-grade serous carcinoma is the most common form of not only ovarian cancer, but also of fallopian tube and peritoneal cancers - and their histological similarities and close anatomical locations have not only led several professional organisations to recommend similar clinical management,^{49,231} but also to hypothesise that they may have similar origins, arising from precursor lesions in the fallopian tubes.¹⁸⁹ The observation that tubal ligation is associated with reductions in risk of all three tumour types would be consistent with this hypothesis.

Furthermore, the finding that tubal ligation was not associated with a reduction in the risk of other hormonally-driven cancers would be consistent with the hypothesis that tubal ligation acts to reduce the risk of these cancer types not primarily by modifying hormone levels, as some have suggested,^{209–211} but is more likely to be related to physically blocking the passage of pro-cancerous agents (whether cells from the endometrium or fallopian tubes, or other agents).

10.3.3 Reproductive factors and risk of ovarian cancer histotypes

The analyses in Chapter 7 confirmed that childbearing is associated with a reduced risk of ovarian cancer overall, with about a 20% reduction in risk associated with the first birth, and a 6% reduction in risk per subsequent birth. There was again significant heterogeneity in the associations with parity by histological tumour type: nulliparous women had a 50-70% increased risk of endometrioid or clear cell tumours compared to women with one child, but smaller-magnitude and/or non-significant increases in risk of mucinous or serous tumours. Similarly, each additional birth was associated with a 25% reduction in risk of clear cell tumours, a 12% reduction in risk of endometrioid tumours, and smaller or non-significant reductions in risk of mucinous and serous tumours.

These findings are consistent with the new hypotheses of ovarian cancer origins: the substantial increase in risk of endometrioid and clear cell tumours amongst nulliparous women could be consistent with a condition (such as endometriosis) that both increases the risk of infertility/ nulliparity, and also increases the risk of these specific histotypes of ovarian cancer.

Breastfeeding was associated with a small but significant reduction in risk of ovarian cancer overall, at least for longer durations: there was about a 10% reduction in risk per 12-months of breastfeeding. This association did not show heterogeneity by histological type, but it is possible that the analysis was under-powered.

10.3.4 Image analysis of the fallopian tubes

Chapter 8 outlined the development of a novel algorithm ('macro') for the identification of the fallopian tube epithelium, to aid image analysis as part of studies of potential novel markers of pre-cancerous changes within the fallopian tubes, and described a series of validation studies that were performed. The results showed that the fallopian tubal

epithelium can be identified in an automated manner with a reasonable degree of reliability, and in a much shorter time, compared to traditional methods of annotating images by hand.

Although there are some systematic differences between the estimates for percentage staining obtained when the epithelium is identified by the macro versus when it is identified by annotation of the image by hand (the gold standard), these differences are comparable to the differences between two human observers in estimates of hand-counting nuclei.

10.3.5 MWS Ovarian cancer tissue pilot study

Chapter 9 described an initial pilot feasibility study of identifying, retrieving and processing histopathology reports and archival tissue samples from participants in the Million Women Study who were treated for ovarian cancer in Oxford.

The first phase of the pilot study has shown that participants who have been treated in Oxford can be identified using hospital records, and that we can link to the databases of the Oxford histopathology services to obtain electronic copies of the pathology reports. Review of the pathology reports has allowed for validation of the routine data on cancer diagnoses we receive from the NHS Central Register and from hospital records.

In the planned second phase of this pilot study, it was intended that tissue samples should be retrieved and processed (scanning selected slides, cutting unstained sections for immunohistochemistry, and taking tissue cores both for constructing tissue microarrays and for storage for possible future DNA extraction). It has unfortunately not been possible to complete this second phase of the pilot study within the timescale of this DPhil, partly due to unexpected capacity constraints at OCHRe. However, it is hoped that work can continue on this project beyond the confines of this DPhil.

10.4 FINDINGS IN CONTEXT: THE NEW HYPOTHESES OF OVARIAN CANCER ORIGINS

The main epidemiological findings for different histotypes of ovarian cancer presented in this thesis (Chapters 5 to 7) constitute an important contribution to the literature on the aetiology of the disease.

The increasing evidence of the potentially disparate origins of the different ovarian cancer histotypes has led to growing interest in examining heterogeneity in epidemiological associations. Given the importance of having sufficient cases of the less-common histotypes, few individual studies can explore these issues, and so the main results of interest have come from collaborative consortia-based individual-patient meta-analyses. The largest consortium to report on tubal ligation and reproductive factors and ovarian cancer risk (the Ovarian Cancer Association Consortium) includes about 10,000 cases of ovarian cancer,^{1,26} but is based on retrospective reporting of exposures, which can be susceptible to considerable bias. There is also a risk of recruitment bias, as women with ovarian cancer are only recruited to retrospective studies after diagnosis of cancer, and thus there is a risk that women with more rapidly-fatal or advanced disease may be less likely to participate.

The largest consortium of prospective studies (the Ovarian Cancer Cohort Consortium) has recently published a summary of ovarian cancer risk factors by histologic subtype within their consortium;²⁷ however, with about 4,000 cases from 13 studies with information on tubal ligation, they still have substantially less statistical power than the Million Women Study to explore heterogeneity by histotype, which is reflected in the fact that for several exposures they have suggestive but non-significant associations, where MWS has clear and significant findings. The results reported here, showing associations by ovarian cancer histotype for the largest prospective study (with more cases than 13

other prospective studies combined in a consortium), are thus an important contribution to this growing literature.

It is particularly important to have reliable epidemiological evidence of histotype-specific associations for ovarian cancer at this time, to aid in evaluating the plausibility of new hypotheses regarding ovarian cancer origins. Several professional organisations and individual authors have argued that the evidence for the new hypotheses (particularly as regards the putative tubal origins for many cases of high-grade serous ovarian cancer) is so convincing that one should act on it now for ovarian cancer prevention. In particular, it has been suggested that the fallopian tubes should be removed in any woman who is having pelvic or abdominal surgery, and who does not want more children, in order to reduce her risk of ovarian cancer ('opportunistic prophylactic salpingectomy').^{232,233}

Some authors have suggested that salpingectomy alone might also be a useful interim step for women with strong genetic risk of ovarian cancer (e.g. carriers of hereditary *BRCA1/2* mutations), if they wish to reduce their risk of ovarian cancer, but do not yet feel ready to have their ovaries removed (oophorectomy).^{234,235} Others, however, have argued that we do not have sufficient evidence to be certain that salpingectomy alone would indeed reduce the risk of ovarian cancer (whether in the general population or in women with *BRCA1/2* mutations),²³⁶ and that this should first be studied in a randomised clinical trial.²³⁷ In addition, for women with *BRCA1/2* mutations, even if salpingectomy reduced the risk of ovarian cancer, it would have no impact on breast cancer (in contrast to oophorectomy, which substantially reduces the risk of breast cancer¹⁵). Non-randomised prospective studies of early salpingectomy with delayed oophorectomy in *BRCA1/2* mutation carriers are currently in progress in the USA (NCT01907789) and in the Netherlands²³⁸ (NCT02321228).

While the epidemiological associations reported here cannot confirm definitively the new hypotheses (e.g. of origins in the fallopian tubes or endometriosis), they can

be evaluated for overall consistency. In general, the results reported here are consistent with the origins hypotheses: tubal ligation is associated with reduced risks of high-grade serous ovarian cancer, and of cancers of the peritoneum and fallopian tubes, all of which are hypothesised to arise from the tubal epithelium. For endometrioid and clear cell ovarian cancers (which are hypothesised to arise from endometriosis), there is a substantial reduction in risk with tubal ligation (possibly due to blocking endometrial cells from the uterus), and a substantial increase in risk with nulliparity (possibly related to endometriosis causing reduced fertility).

Nevertheless, it is important to highlight some findings that might be regarded as less consistent with the new hypotheses, and questions which remain unanswered - particularly as regards tubal ligation and the putative mechanisms behind the associations observed. Firstly, why should the risk of high-grade serous cancer (which is hypothesised to arise from the fimbrial end of the fallopian tube, adjacent to the ovary) be reduced by tubal ligation (which tends to involve the central portion of the tube). Secondly, why should tubal ligation (which is performed on average at age 35 in this cohort) affect the risk of endometriosis-associated endometrioid and clear cell tumours, given that endometriosis is believed to be already established and symptomatic at a younger age.^{239,240} Thirdly, if the association between tubal ligation and ovarian cancer risk is causal and related to a physical blocking effect (e.g. blocking the passage of certain cell types through the fallopian tubes), why is there not evidence of an effect of timing of tubal ligation, with a greater reduction in risk for women who had surgery at a younger age?

Chapters 8 and 9 describe the development of resources for future tissue-based studies of ovarian cancer, to expand further understanding of the origins and characteristics of the different histotypes.

Work done in collaboration with others has already shown that specific markers can

be identified in the fallopian tube epithelium that appear to be associated with high-grade serous ovarian cancer (particularly increased expression of SOX2),²²⁴ and there is ongoing work to further investigate this and other markers. For this continuing research, by both the research groups to which I am affiliated and others, it would be invaluable to have an automated means of identifying the tubal epithelium, to speed image analysis - and thus the macro that has been developed and validated as part of this thesis (Chapter 8) is a useful contribution to this small but growing area of research, and has the potential for significant time-saving.

The pilot study of obtaining pathology reports and archived tissue samples from Million Women Study participants with ovarian cancer (Chapter 9) has helped to establish procedures for possible future expansion of the pilot study to other centres within the UK where MWS participants have been treated for ovarian cancer. The initial purpose of reviewing reports and slides is to confirm the diagnosis of histological type, or to assign a clearer one for cases for which it was uncertain (hopefully thus reducing the ‘other/unspecified’ category). Nevertheless, the scope for immunohistochemical characterisation of the tumours and fallopian tubes, and possibly targeted sequencing in the future, have the potential to make this a valuable resource for ongoing and future research.

10.5 STRENGTHS AND LIMITATIONS OF THE STUDY

The epidemiological analyses conducted in the Million Women Study benefit from the many strengths of this study, including the prospective design, systematic follow-up with nationwide coverage using routine linked data from multiple sources (with less than 1% of the cohort lost to follow-up after more than 14 years), large sample size, and detailed information on multiple potential confounding factors.

The large sample size not only makes possible analyses by ovarian cancer histotype,

but also enables analyses of other uncommon but related cancers (such as cancers of the peritoneum and fallopian tube in Chapter 6). The analyses of other uncommon cancer types was also facilitated by the prospective design, which allows one to look at risk factors for multiple different outcomes.

However, although the analyses presented here represent the largest prospective single study of which I am aware, the cohort is still relatively under-powered for subgroup analyses by histotype (e.g. timing of births, or timing of tubal ligation, by histotype). Thus, while the overall findings should be reliable, there is still a potential that some significant associations in some subgroups may have arisen by chance, or correspondingly that some real associations were not statistically significant here.

One important limitation of this study was that it made use of routinely-collected information on cancer diagnoses and histological types, and did not standardly involve central histopathological review (whether of reports or slides). Not only is there a risk that some tumours might have been mis-coded, but classification systems have also changed over time (e.g. serous carcinomas have changed from a three-tier to a two-tier grading system), and my attempts to ‘convert’ diagnoses from the old to the new systems are likely to be imperfect. However, while it is likely that some tumours have been misclassified, any misclassification would tend to push tests for heterogeneity by histotype towards the null. Thus, the fact that substantial heterogeneity by histotype was found for many different exposures, is additional evidence for the differences between tumour types being genuine (but probably somewhat underestimated here).

More generally, the fact that outcome data (diagnosis of ovarian cancer histotypes) came from routine data sources (e.g. cancer registries) has both strengths and limitations. The obvious strength is that these diagnoses are not dependent on patient recall, and so should be less susceptible to bias. The limitation, however, is that one is instead dependent on routine coding systems; these will no doubt have their own flaws, but the

misclassification is arguably likely to be more at random, and hence introduce random error rather than bias.

Another potential limitation of the epidemiological analyses was that information on both the primary exposures, and potential confounders, was all taken from the recruitment questionnaire, and exposure information was not updated based on resurvey questionnaires. The rationale for this decision was that the primary exposures (e.g. tubal ligation, parity and breastfeeding), and the most important confounder (use of the oral contraceptive pill) would not change in a post-menopausal cohort. While exposure to some other factors might have changed over time (e.g. smoking, alcohol, or use of menopausal hormones), adjustment for these factors made relatively little difference to the main estimates of effect for the primary exposures of interest, and thus changes in exposure with time would also be likely to have little effect.

The validity of the self-reported information in the Million Women Study has been investigated by several previous validation studies. For example, one validation study was conducted amongst 541 women who were participants in both the Million Women Study and the MRC National Survey of Health and Development (NSHD, a cohort study that recruited participants at birth in 1946), and compared self-reported information from the MWS questionnaire(s) to information reported or measured closer to the relevant time from NSHD. This study found a good or excellent level of agreement for many variables relating to body size and reproductive factors, including body mass index and ever-use of the oral contraceptive pill.²⁴¹

Another validation study compared repeated self-reported values for anthropometric variables from MWS questionnaires to clinical measurements at the end of the study period for a sample of 4,000 MWS participants.²⁴² This study found that, in general, the differences between self-reported and measured values were small, and that the values did not change substantially over time.²⁴²

These validation studies thus suggest that the use of self-reported information from the Million Women Study is reasonable, as there is a good level of agreement with measured values, and that updating with later reported measures would probably have made relatively little difference (especially given that the main variables of interest would not have changed in this cohort).

The project to validate a macro for identifying the fallopian tube epithelium had a limitation in its relatively small sample size, with 40 cases in the first phase of the validation study, and 17 in the second phase. However, it had the strength of being able to compare staining using several different methods to identify the epithelium, including hand-annotation, the new macro, and counting of individual nuclei by hand.

The pilot study to obtain archival tissue from Million Women Study participants with ovarian cancer has unfortunately thus far been unable to progress to retrieving tissue samples, partly for reasons of timings and logistics. However, the completed first phase has helped to establish that secure linkage and identification of suitable samples is possible, has allowed for validation of the diagnoses of ovarian cancer obtained from routine data sources, and can inform subsequent planning.

10.6 RECOMMENDATIONS FOR FUTURE RESEARCH

There are several areas for future research that would arise naturally from the work presented in this thesis.

Firstly, there remain several lifestyle or reproductive factors that have been reported to be possibly associated with ovarian cancer risk in other studies, and for which information is available in the Million Women Study, but which have not previously been published on by the Collaborative Group on Epidemiological Studies of Ovarian Cancer, and which have not been investigated in this study. These would include factors such

as age at menarche, age at menopause, and hysterectomy - for all of which information is available from the recruitment questionnaire. Information on endometriosis from the recruitment questionnaire is limited (there was no specific question, but some women reported endometriosis in a free-text section), but additional information on hospital diagnoses of endometriosis is available from hospital records. However, careful analysis would be needed, in order to limit potential biases.

Of particular interest would be future analyses of the risk of ovarian cancer associated with use of common analgesics, and of talcum powder. However, information on both of these exposures was only collected for the first time in resurvey questionnaires, and thus there is currently insufficient follow-up for these analyses to be reliable. Common over-the-counter analgesics such as aspirin and other non-steroidal anti-inflammatory drugs (NSAIDS, e.g. ibuprofen) have been reported by some to be associated with a reduced risk of ovarian cancer.^{137,138} Use of talcum powder (talc) in the perineal/ genital area has been reported by several retrospective studies to be associated with an increased risk of ovarian cancer,¹³³ and has led to several high-profile court cases with large awards of damages.^{243,244}

Secondly, further work remains to be done on pooling all the available published and unpublished data to explore associations between these factors and risk of ovarian cancer by histological type - both for the exposures examined in this thesis (e.g. tubal ligation, parity and breastfeeding), and for other factors of interest. An overview of the ways in which the epidemiological associations vary for the different ovarian cancer histotypes could be invaluable in the ongoing work of trying to determine their origins.

Thirdly, depending on the results of the ongoing pilot study of obtaining tissue samples from MWS participants with ovarian cancer treated in Oxford, there is a wide range of potential future work - both in terms of using the samples from the Oxford pilot study, and expanding the project to encompass women treated at other sites. For

example, the samples obtained could be used to investigate specific immunohistochemical markers of tumour histotype (using tumour TMAs), or of changes in the fallopian tubes (using sections of the fallopian tubes). Depending on developments in methods for extracting DNA from formalin-fixed, paraffin-embedded (FFPE) tissue, it may also be possible in the future to conduct targeted sequencing; whole-genome sequencing of such samples will probably be unfeasible.

Taken together, this work could greatly expand our understanding of the molecular/ genetic, histological, and epidemiological characteristics of the different ovarian cancer subtypes, which would in turn have an impact on the prevention, detection, and treatment of ovarian cancer.

10.7 CONCLUSION

In conclusion, there is convincing evidence that tubal ligation is associated with reduced risks of cancers of the ovary, peritoneum, and fallopian tube, and that childbearing and breastfeeding are associated with reduced risks of ovarian cancer. These associations with ovarian cancer are highly heterogeneous between the different histological tumour types, and these variations are consistent with hypotheses that they may have distinct origins. In particular, risks of endometrioid and clear cell tumours (which are hypothesised to arise from endometriosis) are substantially reduced amongst women with tubal ligation, but are substantially increased amongst nulliparous women. For high-grade serous carcinomas (the most common type of ovarian cancer, and also of fallopian tube and peritoneal cancers), there is a smaller but significant reduction in risk with tubal ligation, but associations with parity were not significant.

This thesis has demonstrated the considerable potential for research at the junction between histopathology and epidemiology, particularly in the field of ovarian cancer, in

which studies of the variation of population-level associations by histotype can help with the evaluation of novel hypotheses, and so expand our understanding of the underlying biology. This thesis has also begun some of the future work that will be needed in this field, developing techniques in image analysis, and the availability of linked biological specimens.

REFERENCES

- [1] W. Sieh, S. Salvador, V. McGuire, R. P. Weber, K. L. Terry, M. A. Rossing, H. Risch, A. H. Wu, P. M. Webb, K. Moysich, J. A. Doherty, A. Felberg, D. Miller, S. J. Jordan, M. T. Goodman, G. Lurie, J. Chang-Claude, A. Rudolph, S. K. Kjaer, A. Jensen, E. Hogdall, E. V. Bandera, S. H. Olson, M. G. King, L. Rodriguez-Rodriguez, L. A. Kiemeny, T. Marees, L. F. Massuger, A. M. van Altena, R. B. Ness, D. W. Cramer, M. C. Pike, C. L. Pearce, A. Berchuck, J. M. Schildkraut, and A. S. Whittemore. Tubal ligation and risk of ovarian cancer subtypes: a pooled analysis of case-control studies. *Int J Epidemiol*, 42(2):579–89, 2013.
- [2] M. Lunn and D. McNeil. Applying Cox regression to competing risks. *Biometrics*, 51(2):524–32, 1995.
- [3] N. E. Allen, V. Beral, D. Casabonne, S. W. Kan, G. K. Reeves, A. Brown, J. Green, and Million Women Study Collaborators. Moderate alcohol intake and cancer incidence in women. *J Natl Cancer Inst*, 101(5):296–305, 2009.
- [4] G. K. Reeves, K. Pirie, V. Beral, J. Green, E. Spencer, D. Bull, and Million Women Study Collaboration. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ*, 335(7630):1134, 2007.
- [5] K. Coffey, V. Beral, J. Green, G. Reeves, and I. Barnes. Lifestyle and reproductive risk factors associated with anal cancer in women aged over 50 years. *Br J Cancer*, 112(9):1568–74, 2015.
- [6] J Ferlay, I Soerjomataram, M Ervik, R Dikshit, S Eser, C Mathers, M Rebelo, DM Parkin, D Forman, and F Bray. GLOBOCAN 2012 v1.0, Cancer incidence and mortality worldwide: IARC CancerBase no. 11 [internet], 2013. Available from: <http://globocan.iarc.fr>, accessed on 06/July/2016.
- [7] M. Sant, M. D. Chirlaque Lopez, R. Agresti, M. J. Sanchez Perez, B. Holleczeck, M. Bielska-Lasota, N. Dimitrova, K. Innos, A. Katalinic, H. Langseth, N. Larranaga, S. Rossi, S. Siesling, P. Minicozzi, and Eurocare-Working Group. Survival of women with cancers of breast and genital organs in Europe 1999-2007: results of the EUROCARE-5 study. *Eur J Cancer*, 2015.
- [8] N Bannister and J Broggio. *Cancer survival by stage at diagnosis for England (experimental statistics): Adults diagnosed 2012, 2013 and 2014 and followed up to 2015*. Statistical Bulletin. Office for National Statistics and Public Health England, 2016.
- [9] S. S. Buys, E. Partridge, M. H. Greene, P. C. Prorok, D. Reding, T. L. Riley, P. Hartge, R. M. Fagerstrom, L. R. Ragard, D. Chia, G. Izmirlian, M. Fouad, C. C. Johnson, J. K. Gohagan, and PLCO Project Team. Ovarian cancer screening in the Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial: findings from the initial screen of a randomized trial. *Am J Obstet Gynecol*, 193(5):1630–9, 2005.

-
- [10] S. S. Buys, E. Partridge, A. Black, C. C. Johnson, L. Lamerato, C. Isaacs, D. J. Reding, R. T. Greenlee, L. A. Yokochi, B. Kessel, E. D. Crawford, T. R. Church, G. L. Andriole, J. L. Weissfeld, M. N. Fouad, D. Chia, B. O'Brien, L. R. Ragard, J. D. Clapp, J. M. Rathmell, T. L. Riley, P. Hartge, P. F. Pinsky, C. S. Zhu, G. Izmirlian, B. S. Kramer, A. B. Miller, J. L. Xu, P. C. Prorok, J. K. Gohagan, C. D. Berg, and PLCO Project Team. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening randomized controlled trial. *JAMA*, 305(22):2295–303, 2011.
 - [11] I. J. Jacobs, U. Menon, A. Ryan, A. Gentry-Maharaj, M. Burnell, J. K. Kalsi, N. N. Amso, S. Apostolidou, E. Benjamin, D. Cruickshank, D. N. Crump, S. K. Davies, A. Dawney, S. Dobbs, G. Fletcher, J. Ford, K. Godfrey, R. Gunu, M. Habib, R. Hallett, J. Herod, H. Jenkins, C. Karpinskyj, S. Leeson, S. J. Lewis, W. R. Liston, A. Lopes, T. Mould, J. Murdoch, D. Oram, D. J. Rabideau, K. Reynolds, I. Scott, M. W. Seif, A. Sharma, N. Singh, J. Taylor, F. Warburton, M. Widschwendter, K. Williamson, R. Woolas, L. Fallowfield, A. J. McGuire, S. Campbell, M. Parmar, and S. J. Skates. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet*, 387(10022):945–56, 2016.
 - [12] J. M. Hall, M. K. Lee, B. Newman, J. E. Morrow, L. A. Anderson, B. Huey, and M. C. King. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science*, 250(4988):1684–9, 1990.
 - [13] Y. Miki, J. Swensen, D. Shattuck-Eidens, P. A. Futreal, K. Harshman, S. Tavtigian, Q. Liu, C. Cochran, L. M. Bennett, W. Ding, R. Bell, J. Rosenthal, C. Hussey, T. Tran, M. McClure, C. Frye, T. Hattier, R. Phelps, A. Haugen-Strano, H. Katcher, K. Yakumo, Z. Gholami, D. Shaffer, S. Stone, S. Bayer, C. Wray, R. Bogden, P. Dayananth, J. Ward, P. Tonin, S. Narod, P. K. Bristow, F. H. Norris, L. Helvering, P. Morrison, P. Rostek, M. Lai, J. C. Barrett, C. Lewis, S. Neuhausen, L. Cannon-Albright, D. Goldgar, R. Wiseman, A. Kamb, and M. H. Skolnick. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, 266(5182):66–71, 1994.
 - [14] R. Wooster, S. L. Neuhausen, J. Mangion, Y. Quirk, D. Ford, N. Collins, K. Nguyen, S. Seal, T. Tran, D. Averill, P. Fields, G. Marshall, S. Narod, G. M. Lenoir, H. Lynch, J. Feunteun, P. Devilee, C. J. Cornelisse, F. H. Menko, P. A. Daly, W. Ormiston, R. McManus, C. Pye, C. M. Lewis, L. A. Cannon-Albright, J. Peto, B. A. J. Ponder, M. H. Skolnick, D. F. Easton, D. E. Goldgar, and M. R. Stratton. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science*, 265(5181):2088–90, 1994.
 - [15] L. C. Hartmann and N. M. Lindor. Risk-reducing surgery in hereditary breast and ovarian cancer. *N Engl J Med*, 374(24):2404, 2016.
 - [16] K. B. Kuchenbaecker, S. J. Ramus, J. Tyrer, A. Lee, H. C. Shen, J. Beesley, K. Lawrenson, L. McGuffog, S. Healey, J. M. Lee, T. J. Spindler, Y. G. Lin, T. Pejovic, Y. Bean, Q. Li, S. Coetzee, D. Hazelett, A. Miron, M. Southey, M. B. Terry, D. E. Goldgar, S. S. Buys, R. Janavicius, C. M. Dorfling, E. J. van Rensburg, S. L. Neuhausen, Y. C. Ding, T. V. Hansen, L. Jonson, A. M. Gerdes, B. Ejlersen, D. Barrowdale, J. Dennis, J. Benitez, A. Osorio, M. J. Garcia, I. Komenaka, J. N. Weitzel, P. Ganschow, P. Peterlongo, L. Bernard, A. Viel, B. Bonanni, B. Peissel, S. Manoukian, P. Radice, L. Papi, L. Ottini, F. Fostira, I. Konstantopoulou, J. Garber, D. Frost, J. Perkins, R. Platte, S. Ellis, Embrace, A. K. Godwin, R. K. Schmutzler, A. Meindl, C. Engel, C. Sutter, O. M. Sinilnikova, Gemo Study Collaborators, F. Damiola, S. Mazoyer, D. Stoppa-Lyonnet, K. Claes, K. De Leeneer, J. Kirk, G. C. Rodriguez, M. Piedmonte, D. M. O'Malley, M. de la Hoya, T. Caldes, K. Aittomaki, H. Nevanlinna, J. M. Collee, M. A. Rookus, J. C. Oosterwijk, Registry Breast Cancer Family, L. Tihomirova, N. Tung, U. Hamann, C. Isacacs, M. Tischkowitz, E. N. Imyanitov, M. A. Caligo, I. G. Campbell, F. B. Hogervorst, Hebon, E. Olah, O. Diez, I. Blanco, J. Brunet, C. Lazaro, M. A. Pujana, A. Jakubowska, J. Gronwald, J. Lubinski, G. Sukiennicki,

REFERENCES

- et al. Identification of six new susceptibility loci for invasive epithelial ovarian cancer. *Nat Genet*, 47(2):164–71, 2015.
- [17] Collaborative Group on Epidemiological Studies of Ovarian Cancer. Ovarian cancer and body size: individual participant meta-analysis including 25,157 women with ovarian cancer from 47 epidemiological studies. *PLoS Med*, 9(4):e1001200, 2012.
- [18] Collaborative Group On Epidemiological Studies Of Ovarian Cancer. Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. *Lancet*, 2015.
- [19] V. Beral, R. Doll, C. Hermon, R. Peto, and G. Reeves. Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies including 23,257 women with ovarian cancer and 87,303 controls. *Lancet*, 371(9609):303–14, 2008.
- [20] K. K. Tsilidis, N. E. Allen, T. J. Key, L. Dossus, A. Lukanova, K. Bakken, E. Lund, A. Fournier, K. Overvad, L. Hansen, A. Tjønneland, V. Fedirko, S. Rinaldi, I. Romieu, F. Clavel-Chapelon, P. Engel, R. Kaaks, M. Schutze, A. Steffen, C. Bamia, A. Trichopoulou, D. Zylis, G. Masala, V. Pala, R. Galasso, R. Tumino, C. Sacerdote, H. B. Bueno-de Mesquita, F. J. van Duijnhoven, M. G. Braem, N. C. Onland-Moret, I. T. Gram, L. Rodriguez, N. Travier, M. J. Sanchez, J. M. Huerta, E. Ardanaz, N. Larranaga, K. Jirstrom, J. Manjer, A. Idahl, N. Ohlson, K. T. Khaw, N. Wareham, T. Mouw, T. Norat, and E. Riboli. Oral contraceptive use and reproductive factors and risk of ovarian cancer in the European Prospective Investigation into Cancer and Nutrition. *Br J Cancer*, 105(9):1436–42, 2011.
- [21] R. J. Kurman and Ie M Shih. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer—shifting the paradigm. *Hum Pathol*, 42(7):918–31, 2011.
- [22] J. Prat. Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features. *Virchows Arch*, 460(3):237–49, 2012.
- [23] R. J. Kurman, M. L. Carcangiu, C. S. Herrington, and R. H. Young (Eds.). *WHO Classification of Tumours of Female Reproductive Organs*. IARC, Lyon, 4th edition, 2014.
- [24] M. Kobel, S. E. Kalloger, N. Boyd, S. McKinney, E. Mehl, C. Palmer, S. Leung, N. J. Bowen, D. N. Ionescu, A. Rajput, L. M. Prentice, D. Miller, J. Santos, K. Swenerton, C. B. Gilks, and D. Huntsman. Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. *PLoS Med*, 5(12):e232, 2008.
- [25] W. G. McCluggage. Morphological subtypes of ovarian carcinoma: a review with emphasis on new developments and pathogenesis. *Pathology*, 43(5):420–32, 2011.
- [26] C. L. Pearce, M. A. Rossing, A. W. Lee, R. B. Ness, P. M. Webb, G. Chenevix-Trench, S. M. Jordan, D. A. Stram, J. Chang-Claude, R. Hein, S. Nickels, G. Lurie, P. J. Thompson, M. E. Carney, M. T. Goodman, K. Moysich, E. Hogdall, A. Jensen, E. L. Goode, B. L. Fridley, J. M. Cunningham, R. A. Vierkant, R. P. Weber, A. Ziogas, H. Anton-Culver, S. A. Gayther, A. Gentry-Maharaj, U. Menon, S. J. Ramus, L. Brinton, N. Wentzensen, J. Lissowska, M. Garcia-Closas, L. F. Massuger, L. A. Kiemeny, A. M. Van Altena, K. K. Aben, A. Berchuck, J. A. Doherty, E. Iversen, V. McGuire, P. G. Moorman, P. Pharoah, M. C. Pike, H. Risch, W. Sieh, D. O. Stram, K. L. Terry, A. Whittemore, A. H. Wu, J. M. Schildkraut, and S. K. Kjaer. Combined and interactive effects of environmental and GWAS-identified risk factors in ovarian cancer. *Cancer Epidemiol Biomarkers Prev*, 22(5):880–90, 2013.
- [27] N. Wentzensen, E. M. Poole, B. Trabert, E. White, A. A. Arslan, A. V. Patel, V. W. Setiawan, K. Visvanathan, E. Weiderpass, H. O. Adami, A. Black, L. Bernstein, L. A. Brinton, J. Buring, L. M. Butler, S. Chamosa, T. V. Clendenen, L. Dossus, R. Fortner, S. M. Gapstur, M. M. Gaudet,

- I. T. Gram, P. Hartge, J. Hoffman-Bolton, A. Idahl, M. Jones, R. Kaaks, V. Kirsh, W. P. Koh, J. V. Lacey, Jr., I. M. Lee, E. Lundin, M. A. Merritt, N. C. Onland-Moret, U. Peters, J. N. Poynter, S. Rinaldi, K. Robien, T. Rohan, D. P. Sandler, C. Schairer, L. J. Schouten, L. K. Sjöholm, S. Sieri, A. Swerdlow, A. Tjønneland, R. Travis, A. Trichopoulou, P. A. van den Brandt, L. Wilkens, A. Wolk, H. P. Yang, A. Zeleniuch-Jacquotte, and S. S. Tworoger. Ovarian cancer risk factors by histologic subtype: An analysis from the Ovarian Cancer Cohort Consortium. *J Clin Oncol*, 2016.
- [28] K. Ohkawa, H. Amasaki, Y. Terashima, S. Aizawa, and E. Ishikawa. Clear cell carcinoma of the ovary: light and electron microscopic studies. *Cancer*, 40(6):3019–29, 1977.
- [29] J. Rosai, USCAP (United States and Canadian Academy of Pathology), and Aperio (Leica Biosystems). Low grade papillary serous ovarian carcinoma. Seminar 1452, Case 18, ID: 18303. *Juan Rosai's collection of surgical pathology seminars*, Available from: <http://www.rosaicollection.org/casedetails.cfm?c=18303>, Accessed 20th Sept 2016.
- [30] J. Rosai, USCAP (United States and Canadian Academy of Pathology), and Aperio (Leica Biosystems). High grade serous ovarian carcinoma. Seminar 1161, Case 8, ID: 14718. *Juan Rosai's collection of surgical pathology seminars*, Available from: <http://www.rosaicollection.org/casedetails.cfm?c=14950>, Accessed 20th Sept 2016.
- [31] J. Rosai, USCAP (United States and Canadian Academy of Pathology), and Aperio (Leica Biosystems). Endometrioid ovarian carcinoma (low grade). Seminar 1184, Case 8, ID: 14950. *Juan Rosai's collection of surgical pathology seminars*, Available from: <http://www.rosaicollection.org/casedetails.cfm?c=14950>, Accessed 20th Sept 2016.
- [32] J. Rosai, USCAP (United States and Canadian Academy of Pathology), and Aperio (Leica Biosystems). Clear cell ovarian cancer. Seminar 1024, Case 16, ID: 2908. *Juan Rosai's collection of surgical pathology seminars*, Available from: <http://www.rosaicollection.org/casedetails.cfm?c=2908>, Accessed 20th Sept 2016.
- [33] J. Rosai, USCAP (United States and Canadian Academy of Pathology), and Aperio (Leica Biosystems). Mucinous ovarian carcinoma. Seminar 692, Case 4, ID: 8979. *Juan Rosai's collection of surgical pathology seminars*, Available from: <http://www.rosaicollection.org/casedetails.cfm?c=8979>, Accessed 20th Sept 2016.
- [34] J. M. Piek, P. J. van Diest, R. P. Zweemer, J. W. Jansen, R. J. Poort-Keesom, F. H. Menko, J. J. Gille, A. P. Jongsma, G. Pals, P. Kenemans, and R. H. Verheijen. Dysplastic changes in prophylactically removed fallopian tubes of women predisposed to developing ovarian cancer. *J Pathol*, 195(4):451–6, 2001.
- [35] F. Medeiros, M. G. Muto, Y. Lee, J. A. Elvin, M. J. Callahan, C. Feltmate, J. E. Garber, D. W. Cramer, and C. P. Crum. The tubal fimbria is a preferred site for early adenocarcinoma in women with familial ovarian cancer syndrome. *Am J Surg Pathol*, 30(2):230–6, 2006.
- [36] M. J. Callahan, C. P. Crum, F. Medeiros, D. W. Kindelberger, J. A. Elvin, J. E. Garber, C. M. Feltmate, R. S. Berkowitz, and M. G. Muto. Primary fallopian tube malignancies in BRCA-positive women undergoing surgery for ovarian cancer risk reduction. *J Clin Oncol*, 25(25):3985–90, 2007.
- [37] A. K. Folkins, E. A. Jarboe, A. Saleemuddin, Y. Lee, M. J. Callahan, R. Drapkin, J. E. Garber, M. G. Muto, S. Tworoger, and C. P. Crum. A candidate precursor to pelvic serous cancer (p53 signature) and its prevalence in ovaries and fallopian tubes from women with BRCA mutations. *Gynecol Oncol*, 109(2):168–73, 2008.

REFERENCES

- [38] D. W. Kindelberger, Y. Lee, A. Miron, M. S. Hirsch, C. Feltmate, F. Medeiros, M. J. Callahan, E. O. Garner, R. W. Gordon, C. Birch, R. S. Berkowitz, M. G. Muto, and C. P. Crum. Intraepithelial carcinoma of the fimbria and pelvic serous carcinoma: Evidence for a causal relationship. *Am J Surg Pathol*, 31(2):161–9, 2007.
- [39] M. H. Roh, D. Kindelberger, and C. P. Crum. Serous tubal intraepithelial carcinoma and the dominant ovarian mass: clues to serous tumor origin? *Am J Surg Pathol*, 33(3):376–83, 2009.
- [40] C. G. Przybycin, R. J. Kurman, B. M. Ronnett, Ie M. Shih, and R. Vang. Are all pelvic (nonuterine) serous carcinomas of tubal origin? *Am J Surg Pathol*, 34(10):1407–16, 2010.
- [41] P. M. Diniz, J. P. Carvalho, E. C. Barakat, and F. M. Carvalho. Fallopian tube origin of supposed ovarian high-grade serous carcinomas. *Clinics (Sao Paulo)*, 66(1):73–6, 2011.
- [42] S. Tang, K. Onuma, P. Deb, E. Wang, A. Lytwyn, M. Sur, and D. Daya. Frequency of serous tubal intraepithelial carcinoma in various gynecologic malignancies: a study of 300 consecutive cases. *Int J Gynecol Pathol*, 31(2):103–10, 2012.
- [43] N. Koc, S. Ayas, and L. Uygur. The association of serous tubal intraepithelial carcinoma with gynecologic pathologies and its role in pelvic serous cancer. *Gynecol Oncol*, 134(3):486–91, 2014.
- [44] B. E. Howitt, S. Hanamornroongruang, D. I. Lin, J. E. Conner, S. Schulte, N. Horowitz, C. P. Crum, and E. E. Meserve. Evidence for a dualistic model of high-grade serous carcinoma: BRCA mutation status, histology, and tubal intraepithelial carcinoma. *Am J Surg Pathol*, 39(3):287–93, 2015.
- [45] J. D. Seidman. Serous tubal intraepithelial carcinoma localizes to the tubal-peritoneal junction: a pivotal clue to the site of origin of extrauterine high-grade serous carcinoma (ovarian cancer). *Int J Gynecol Pathol*, 34(2):112–20, 2015.
- [46] S. Munakata and T. Yamamoto. Incidence of serous tubal intraepithelial carcinoma (STIC) by algorithm classification in serous ovarian tumor associated with PAX8 expression in tubal epithelia: a study of single institution in Japan. *Int J Gynecol Pathol*, 34(1):9–18, 2015.
- [47] K. Malmberg, C. Klynning, A. Floter-Radestad, and J. W. Carlson. Serous tubal intraepithelial carcinoma, chronic fallopian tube injury, and serous carcinoma development. *Virchows Arch*, 468(6):707–13, 2016.
- [48] J. S. Berek, C. Crum, and M. Friedlander. Cancer of the ovary, fallopian tube, and peritoneum. *Int J Gynaecol Obstet*, 131 Suppl 2:S111–22, 2015.
- [49] National Comprehensive Cancer Network. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Ovarian cancer including fallopian tube cancer and primary peritoneal cancer, Version 1.2016*. National Comprehensive Cancer Network, Ft. Washington, PA, 2016. Available from: https://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf.
- [50] C. B. Gilks, J. Irving, M. Kobel, C. Lee, N. Singh, N. Wilkinson, and W. G. McCluggage. Incidental nonuterine high-grade serous carcinomas arise in the fallopian tube in most cases: further evidence for the tubal origin of high-grade serous carcinomas. *Am J Surg Pathol*, 39(3):357–64, 2015.
- [51] J. C. Morrison, L. Z. Blanco, Jr., R. Vang, and B. M. Ronnett. Incidental serous tubal intraepithelial carcinoma and early invasive serous carcinoma in the nonprophylactic setting: analysis of a case series. *Am J Surg Pathol*, 39(4):442–53, 2015.
- [52] J. T. Rabban, K. Garg, B. Crawford, L. M. Chen, and C. J. Zaloudek. Early detection of high-grade tubal serous carcinoma in women at low risk for hereditary breast and ovarian cancer syndrome by systematic examination of fallopian tubes incidentally removed during benign surgery. *Am J Surg Pathol*, 38(6):729–42, 2014.

-
- [53] A. A. Ahmed, D. Etemadmoghadam, J. Temple, A. G. Lynch, M. Riad, R. Sharma, C. Stewart, S. Fereday, C. Caldas, A. Defazio, D. Bowtell, and J. D. Brenton. Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J Pathol*, 221(1):49–56, 2010.
- [54] A. Malpica, M. T. Deavers, K. Lu, D. C. Bodurka, E. N. Atkinson, D. M. Gershenson, and E. G. Silva. Grading ovarian serous carcinoma using a two-tier system. *Am J Surg Pathol*, 28(4):496–504, 2004.
- [55] G. Singer, R. Oldt 3rd, Y. Cohen, B. G. Wang, D. Sidransky, R. J. Kurman, and Ie M. Shih. Mutations in BRAF and KRAS characterize the development of low-grade ovarian serous carcinoma. *J Natl Cancer Inst*, 95(6):484–6, 2003.
- [56] G. Singer, R. Stohr, L. Cope, R. Dehari, A. Hartmann, D. F. Cao, T. L. Wang, R. J. Kurman, and Ie M. Shih. Patterns of p53 mutations separate ovarian serous borderline tumors and low- and high-grade carcinomas and provide support for a new model of ovarian carcinogenesis: a mutational analysis with immunohistochemical correlation. *Am J Surg Pathol*, 29(2):218–24, 2005.
- [57] J. Li, N. Abushahin, S. Pang, L. Xiang, S. K. Chambers, O. Fadare, B. Kong, and W. Zheng. Tubal origin of ‘ovarian’ low-grade serous carcinoma. *Mod Pathol*, 24(11):1488–99, 2011.
- [58] R. Vang, Ie M. Shih, and R. J. Kurman. Fallopian tube precursors of ovarian low- and high-grade serous neoplasms. *Histopathology*, 62(1):44–58, 2013.
- [59] J. Prat. New insights into ovarian cancer pathology. *Ann Oncol*, 23 Suppl 10:x111–7, 2012.
- [60] A. LaGrenade and S. G. Silverberg. Ovarian tumors associated with atypical endometriosis. *Hum Pathol*, 19(9):1080–4, 1988.
- [61] P. D. DePriest, E. R. Banks, D. E. Powell, J. R. van Nagell, Jr, H. H. Gallion, L. E. Puls, J. E. Hunter, R. J. Kryscio, and M. B. Royalty. Endometrioid carcinoma of the ovary and endometriosis: the association in postmenopausal women. *Gynecol Oncol*, 47(1):71–5, 1992.
- [62] R. Sainz de la Cuesta, J. H. Eichhorn, L. W. Rice, A. F. Fuller, Jr., N. Nikrui, and B. A. Goff. Histologic transformation of benign endometriosis to early epithelial ovarian cancer. *Gynecol Oncol*, 60(2):238–44, 1996.
- [63] M. Fukunaga, K. Nomura, E. Ishikawa, and S. Ushigome. Ovarian atypical endometriosis: its close association with malignant epithelial tumours. *Histopathology*, 30(3):249–55, 1997.
- [64] S. Ogawa, T. Kaku, S. Amada, H. Kobayashi, T. Hirakawa, K. Ariyoshi, T. Kamura, and H. Nakano. Ovarian endometriosis associated with ovarian carcinoma: a clinicopathological and immunohistochemical study. *Gynecol Oncol*, 77(2):298–304, 2000.
- [65] H. Yoshikawa, H. Jimbo, S. Okada, K. Matsumoto, T. Onda, T. Yasugi, and Y. Taketani. Prevalence of endometriosis in ovarian cancer. *Gynecol Obstet Invest*, 50 Suppl 1:11–7, 2000.
- [66] R. C. Stern, R. Dash, R. C. Bentley, M. J. Snyder, A. F. Haney, and S. J. Robboy. Malignancy in endometriosis: frequency and comparison of ovarian and extraovarian types. *Int J Gynecol Pathol*, 20(2):133–9, 2001.
- [67] H. Jimbo, H. Yoshikawa, T. Onda, T. Yasugi, A. Sakamoto, and Y. Taketani. Prevalence of ovarian endometriosis in epithelial ovarian cancer. *Int J Gynaecol Obstet*, 59(3):245–50, 1997.
- [68] X. Jiang, S. J. Morland, A. Hitchcock, E. J. Thomas, and I. G. Campbell. Allelotyping of endometriosis with adjacent ovarian carcinoma reveals evidence of a common lineage. *Cancer Res*, 58(8):1707–12, 1998.

REFERENCES

- [69] A. H. Prowse, S. Manek, R. Varma, J. Liu, A. K. Godwin, E. R. Maher, I. P. Tomlinson, and S. H. Kennedy. Molecular genetic evidence that endometriosis is a precursor of ovarian cancer. *Int J Cancer*, 119(3):556–62, 2006.
- [70] B. Xu, S. Hamada, I. Kusuki, R. Itoh, and J. Kitawaki. Possible involvement of loss of heterozygosity in malignant transformation of ovarian endometriosis. *Gynecol Oncol*, 120(2):239–46, 2011.
- [71] S. Jones, T. L. Wang, Ie M. Shih, T. L. Mao, K. Nakayama, R. Roden, R. Glas, D. Slamon, L. A. Diaz, Jr., B. Vogelstein, K. W. Kinzler, V. E. Velculescu, and N. Papadopoulos. Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science*, 330(6001):228–31, 2010.
- [72] K. C. Wiegand, S. P. Shah, O. M. Al-Agha, Y. Zhao, K. Tse, T. Zeng, J. Senz, M. K. McConechy, M. S. Anglesio, S. E. Kalloger, W. Yang, A. Heravi-Moussavi, R. Giuliany, C. Chow, J. Fee, A. Zayed, L. Prentice, N. Melnyk, G. Turashvili, A. D. Delaney, J. Madore, S. Yip, A. W. McPherson, G. Ha, L. Bell, S. Fereday, A. Tam, L. Galletta, P. N. Tonin, D. Provencher, D. Miller, S. J. Jones, R. A. Moore, G. B. Morin, A. Oloumi, N. Boyd, S. A. Aparicio, Ie M. Shih, A. M. Mes-Masson, D. D. Bowtell, M. Hirst, B. Gilks, M. A. Marra, and D. G. Huntsman. ARID1A mutations in endometriosis-associated ovarian carcinomas. *N Engl J Med*, 363(16):1532–43, 2010.
- [73] A. Ayhan, T. L. Mao, T. Seckin, C. H. Wu, B. Guan, H. Ogawa, M. Futagami, H. Mizukami, Y. Yokoyama, R. J. Kurman, and Ie M. Shih. Loss of ARID1A expression is an early molecular event in tumor progression from ovarian endometriotic cyst to clear cell and endometrioid carcinoma. *Int J Gynecol Cancer*, 22(8):1310–5, 2012.
- [74] M. K. McConechy, J. Ding, J. Senz, W. Yang, N. Melnyk, A. A. Tone, L. M. Prentice, K. C. Wiegand, J. N. McAlpine, S. P. Shah, C. H. Lee, P. J. Goodfellow, C. B. Gilks, and D. G. Huntsman. Ovarian and endometrial endometrioid carcinomas have distinct CTNNB1 and PTEN mutation profiles. *Mod Pathol*, 27(1):128–34, 2014.
- [75] G. Chene, V. Ouellet, K. Rahimi, V. Barres, D. Provencher, and A. M. Mes-Masson. The ARID1A pathway in ovarian clear cell and endometrioid carcinoma, contiguous endometriosis, and benign endometriosis. *Int J Gynaecol Obstet*, 130(1):27–30, 2015.
- [76] L. Catasus, E. Bussaglia, I. Rodriguez, A. Gallardo, C. Pons, J. A. Irving, and J. Prat. Molecular genetic alterations in endometrioid carcinomas of the ovary: similar frequency of beta-catenin abnormalities but lower rate of microsatellite instability and PTEN alterations than in uterine endometrioid carcinomas. *Hum Pathol*, 35(11):1360–8, 2004.
- [77] K. T. Kuo, T. L. Mao, S. Jones, E. Veras, A. Ayhan, T. L. Wang, R. Glas, D. Slamon, V. E. Velculescu, R. J. Kurman, and Ie M. Shih. Frequent activating mutations of PIK3CA in ovarian clear cell carcinoma. *Am J Pathol*, 174(5):1597–601, 2009.
- [78] S. Yamamoto, H. Tsuda, M. Takano, S. Tamai, and O. Matsubara. Loss of ARID1A protein expression occurs as an early event in ovarian clear-cell carcinoma development and frequently coexists with PIK3CA mutations. *Mod Pathol*, 25(4):615–24, 2012.
- [79] M. A. Riopel, B. M. Ronnett, and R. J. Kurman. Evaluation of diagnostic criteria and behavior of ovarian intestinal-type mucinous tumors: atypical proliferative (borderline) tumors and intraepithelial, microinvasive, invasive, and metastatic carcinomas. *Am J Surg Pathol*, 23(6):617–35, 1999.
- [80] J. D. Seidman and F. Khedmati. Exploring the histogenesis of ovarian mucinous and transitional cell (Brenner) neoplasms and their relationship with Walthard cell nests: a study of 120 tumors. *Arch Pathol Lab Med*, 132(11):1753–60, 2008.

-
- [81] J. D. Seidman, A. Yemelyanova, R. J. Zaino, and R. J. Kurman. The fallopian tube-peritoneal junction: a potential site of carcinogenesis. *Int J Gynecol Pathol*, 30(1):4–11, 2011.
 - [82] R. Vang, A. M. Gown, C. Zhao, T. S. Barry, C. Isacson, M. S. Richardson, and B. M. Ronnett. Ovarian mucinous tumors associated with mature cystic teratomas: morphologic and immuno-histochemical analysis identifies a subset of potential teratomatous origin that shares features of lower gastrointestinal tract mucinous tumors more commonly encountered as secondary tumors in the ovary. *Am J Surg Pathol*, 31(6):854–69, 2007.
 - [83] P. F. Pinsky, C. Zhu, S. J. Skates, A. Black, E. Partridge, S. S. Buys, and C. D. Berg. Potential effect of the risk of ovarian cancer algorithm (ROCA) on the mortality outcome of the Prostate, Lung, Colorectal and Ovarian (PLCO) trial. *Int J Cancer*, 132(9):2127–33, 2013.
 - [84] A. Antoniou, P. D. Pharoah, S. Narod, H. A. Risch, J. E. Eyfjord, J. L. Hopper, N. Loman, H. Olsson, O. Johannsson, A. Borg, B. Pasini, P. Radice, S. Manoukian, D. M. Eccles, N. Tang, E. Olah, H. Anton-Culver, E. Warner, J. Lubinski, J. Gronwald, B. Gorski, H. Tulinius, S. Thorlacius, H. Eerola, H. Nevanlinna, K. Syrjakoski, O. P. Kallioniemi, D. Thompson, C. Evans, J. Peto, F. Lalloo, D. G. Evans, and D. F. Easton. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet*, 72(5):1117–30, 2003.
 - [85] A. C. Antoniou, A. P. Cunningham, J. Peto, D. G. Evans, F. Lalloo, S. A. Narod, H. A. Risch, J. E. Eyfjord, J. L. Hopper, M. C. Southey, H. Olsson, O. Johannsson, A. Borg, B. Pasini, P. Radice, S. Manoukian, D. M. Eccles, N. Tang, E. Olah, H. Anton-Culver, E. Warner, J. Lubinski, J. Gronwald, B. Gorski, L. Tryggvadottir, K. Syrjakoski, O. P. Kallioniemi, H. Eerola, H. Nevanlinna, P. D. Pharoah, and D. F. Easton. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer*, 98(8):1457–66, 2008.
 - [86] N. Mavaddat, D. Barrowdale, I. L. Andrulis, S. M. Domchek, D. Eccles, H. Nevanlinna, S. J. Ramus, A. Spurdle, M. Robson, M. Sherman, A. M. Mulligan, F. J. Couch, C. Engel, L. McGuffog, S. Healey, O. M. Sinilnikova, M. C. Southey, M. B. Terry, D. Goldgar, F. O’Malley, E. M. John, R. Janavicius, L. Tihomirova, T. V. Hansen, F. C. Nielsen, A. Osorio, A. Stavropoulou, J. Benitez, S. Manoukian, B. Peissel, M. Barile, S. Volorio, B. Pasini, R. Dolcetti, A. L. Putignano, L. Ottini, P. Radice, U. Hamann, M. U. Rashid, F. B. Hogervorst, M. Kriege, R. B. van der Luijt, HEBON, S. Peock, D. Frost, D. G. Evans, C. Brewer, L. Walker, M. T. Rogers, L. E. Side, C. Houghton, EMBRACE, J. Weaver, A. K. Godwin, R. K. Schmutzler, B. Wappenschmidt, A. Meindl, K. Kast, N. Arnold, D. Niederacher, C. Sutter, H. Deissler, D. Gadzicki, S. Preisler-Adams, R. Varon-Mateeva, I. Schonbuchner, H. Gevensleben, D. Stoppa-Lyonnet, M. Belotti, L. Barjhoux, GEMO Study Collaborators, C. Isaacs, B. N. Peshkin, T. Caldes, M. de la Hoya, C. Canadas, T. Heikkinen, P. Heikkila, K. Aittomaki, I. Blanco, C. Lazaro, J. Brunet, B. A. Agnars-son, A. Arason, R. B. Barkardottir, M. Dumont, J. Simard, M. Montagna, S. Agata, E. D’Andrea, M. Yan, S. Fox, kConFab Investigators, T. R. Rebbeck, W. Rubinstein, N. Tung, J. E. Garber, X. Wang, Z. Fredericksen, V. S. Pankratz, et al. Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). *Cancer Epidemiol Biomarkers Prev*, 21(1):134–47, 2012.
 - [87] National Comprehensive Cancer Network. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Genetic/familial high-risk assessment: Breast and ovarian*. National Comprehensive Cancer Network, Ft Washington, PA, 2015. Available from: https://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf.
 - [88] A Kulkarni and AF Brady. *Scientific Impact Paper: Management of women with a genetic predisposition to gynaecological cancers*, volume No.48 of *Scientific Impact Paper*. Royal College of Obstetricians & Gynaecologists, London, 2015.

REFERENCES

- [89] J. M. Helder-Woolderink, E. A. Blok, H. F. Vasen, H. Hollema, M. J. Mourits, and G. H. De Bock. Ovarian cancer in Lynch syndrome; a systematic review. *Eur J Cancer*, 55:65–73, 2016.
- [90] K. A. McBride, M. L. Ballinger, E. Killick, J. Kirk, M. H. Tattersall, R. A. Eeles, D. M. Thomas, and G. Mitchell. Li-Fraumeni syndrome: cancer risk assessment and clinical management. *Nat Rev Clin Oncol*, 11(5):260–71, 2014.
- [91] T. Walsh, S. Casadei, M. K. Lee, C. C. Pennil, A. S. Nord, A. M. Thornton, W. Roeb, K. J. Agnew, S. M. Stray, A. Wickramanayake, B. Norquist, K. P. Pennington, R. L. Garcia, M. C. King, and E. M. Swisher. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proc Natl Acad Sci U S A*, 108(44):18032–7, 2011.
- [92] H. Song, S. J. Ramus, J. Tyrer, K. L. Bolton, A. Gentry-Maharaj, E. Wozniak, H. Anton-Culver, J. Chang-Claude, D. W. Cramer, R. DiCioccio, T. Dork, E. L. Goode, M. T. Goodman, J. M. Schildkraut, T. Sellers, L. Baglietto, M. W. Beckmann, J. Beesley, J. Blaakaer, M. E. Carney, S. Chanock, Z. Chen, J. M. Cunningham, E. Dicks, J. A. Doherty, M. Durst, A. B. Ekici, D. Fenstermacher, B. L. Fridley, G. Giles, M. E. Gore, I. De Vivo, P. Hillemanns, C. Hogdall, E. Hogdall, E. S. Iversen, I. J. Jacobs, A. Jakubowska, D. Li, J. Lissowska, J. Lubinski, G. Lurie, V. McGuire, J. McLaughlin, K. Medrek, P. G. Moorman, K. Moysich, S. Narod, C. Phelan, C. Pye, H. Risch, I. B. Runnebaum, G. Severi, M. Southey, D. O. Stram, F. C. Thiel, K. L. Terry, Y. Y. Tsai, S. S. Tworoger, D. J. Van Den Berg, R. A. Vierkant, S. Wang-Gohrke, P. M. Webb, L. R. Wilkens, A. H. Wu, H. Yang, W. Brewster, A. Ziogas, Australian Cancer Study, Australian Ovarian Cancer Study Group, Ovarian Cancer Association Consortium, R. Houlston, I. Tomlinson, A. S. Whittemore, M. A. Rossing, B. A. Ponder, C. L. Pearce, R. B. Ness, U. Menon, S. K. Kjaer, J. Gronwald, M. Garcia-Closas, P. A. Fasching, D. F. Easton, G. Chenevix-Trench, A. Berchuck, P. D. Pharoah, and S. A. Gayther. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. *Nat Genet*, 41(9):996–1000, 2009.
- [93] E. L. Goode, G. Chenevix-Trench, H. Song, S. J. Ramus, M. Notaridou, K. Lawrenson, M. Widschwendter, R. A. Vierkant, M. C. Larson, S. K. Kjaer, M. J. Birrer, A. Berchuck, J. Schildkraut, I. Tomlinson, L. A. Kiemeny, L. S. Cook, J. Gronwald, M. Garcia-Closas, M. E. Gore, I. Campbell, A. S. Whittemore, R. Sutphen, C. Phelan, H. Anton-Culver, C. L. Pearce, D. Lambrechts, M. A. Rossing, J. Chang-Claude, K. B. Moysich, M. T. Goodman, T. Dork, H. Nevanlinna, R. B. Ness, T. Rafnar, C. Hogdall, E. Hogdall, B. L. Fridley, J. M. Cunningham, W. Sieh, V. McGuire, A. K. Godwin, D. W. Cramer, D. Hernandez, D. Levine, K. Lu, E. S. Iversen, R. T. Palmieri, R. Houlston, A. M. van Altena, K. K. Aben, L. F. Massuger, A. Brooks-Wilson, L. E. Kelemen, N. D. Le, A. Jakubowska, J. Lubinski, K. Medrek, A. Stafford, D. F. Easton, J. Tyrer, K. L. Bolton, P. Harington, D. Eccles, A. Chen, A. N. Molina, B. N. Davila, H. Arango, Y. Y. Tsai, Z. Chen, H. A. Risch, J. McLaughlin, S. A. Narod, A. Ziogas, W. Brewster, A. Gentry-Maharaj, U. Menon, A. H. Wu, D. O. Stram, M. C. Pike, Wellcome Trust Case-Control Consortium, J. Beesley, P. M. Webb, Australian Cancer Study (Ovarian Cancer), Australian Ovarian Cancer Study Group, X. Chen, A. B. Ekici, F. C. Thiel, M. W. Beckmann, H. Yang, N. Wentzensen, J. Lissowska, P. A. Fasching, E. Despierre, F. Amant, I. Vergote, J. Doherty, R. Hein, S. Wang-Gohrke, G. Lurie, M. E. Carney, et al. A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. *Nat Genet*, 42(10):874–9, 2010.
- [94] K. L. Bolton, J. Tyrer, H. Song, S. J. Ramus, M. Notaridou, C. Jones, T. Sher, A. Gentry-Maharaj, E. Wozniak, Y. Y. Tsai, J. Weidhaas, D. Paik, D. J. Van Den Berg, D. O. Stram, C. L. Pearce, A. H. Wu, W. Brewster, H. Anton-Culver, A. Ziogas, S. A. Narod, D. A. Levine, S. B. Kaye, R. Brown, J. Paul, J. Flanagan, W. Sieh, V. McGuire, A. S. Whittemore, I. Campbell, M. E. Gore, J. Lissowska, H. P. Yang, K. Medrek, J. Gronwald, J. Lubinski, A. Jakubowska, N. D. Le, L. S. Cook, L. E. Kelemen, A. Brooks-Wilson, L. F. Massuger, L. A. Kiemeny, K. K. Aben, A. M. van Altena, R. Houlston, I. Tomlinson, R. T. Palmieri, P. G. Moorman, J. Schildkraut, E. S. Iversen, C. Phelan, R. A. Vierkant, J. M. Cunningham, E. L. Goode, B. L. Fridley, S. Kruger-Kjaer, J. Blaaker, E. Hogdall, C. Hogdall, J. Gross, B. Y. Karlan, R. B. Ness, R. P. Edwards,

- K. Odunsi, K. B. Moyisch, J. A. Baker, F. Modugno, T. Heikkinen, R. Butzow, H. Nevanlinna, A. Leminen, N. Bogdanova, N. Antonenkova, T. Doerk, P. Hillemanns, M. Durst, I. Runnebaum, P. J. Thompson, M. E. Carney, M. T. Goodman, G. Lurie, S. Wang-Gohrke, R. Hein, J. Chang-Claude, M. A. Rossing, K. L. Cushing-Haugen, J. Doherty, C. Chen, T. Rafnar, S. Besenbacher, P. Sulem, K. Stefansson, M. J. Birrer, K. L. Terry, D. Hernandez, D. W. Cramer, I. Vergote, F. Amant, D. Lambrechts, E. Despierre, et al. Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nat Genet*, 42(10):880–4, 2010.
- [95] P. D. Pharoah, Y. Y. Tsai, S. J. Ramus, C. M. Phelan, E. L. Goode, K. Lawrenson, M. Buckley, B. L. Fridley, J. P. Tyrer, H. Shen, R. Weber, R. Karevan, M. C. Larson, H. Song, D. C. Tessier, F. Bacot, D. Vincent, J. M. Cunningham, J. Dennis, E. Dicks, Australian Cancer Study, Australian Ovarian Cancer Study Group, K. K. Aben, H. Anton-Culver, N. Antonenkova, S. M. Armasu, L. Baglietto, E. V. Bandera, M. W. Beckmann, M. J. Birrer, G. Bloom, N. Bogdanova, J. D. Brenton, L. A. Brinton, A. Brooks-Wilson, R. Brown, R. Butzow, I. Campbell, M. E. Carney, R. S. Carvalho, J. Chang-Claude, Y. A. Chen, Z. Chen, W. H. Chow, M. S. Cicek, G. Coetzee, L. S. Cook, D. W. Cramer, C. Cybulski, A. Dansonka-Mieszkowska, E. Despierre, J. A. Doherty, T. Dork, A. du Bois, M. Durst, D. Eccles, R. Edwards, A. B. Ekici, P. A. Fasching, D. Fenstermacher, J. Flanagan, Y. T. Gao, M. Garcia-Closas, A. Gentry-Maharaj, G. Giles, A. Gjyshi, M. Gore, J. Gronwald, Q. Guo, M. K. Halle, P. Harter, A. Hein, F. Heitz, P. Hillemanns, M. Hoatlin, E. Hogdall, C. K. Hogdall, S. Hosono, A. Jakubowska, A. Jensen, K. R. Kalli, B. Y. Karlan, L. E. Kelemen, L. A. Kiemeny, S. K. Kjaer, G. E. Konecny, C. Krakstad, J. Kupryjanczyk, D. Lambrechts, S. Lambrechts, N. D. Le, N. Lee, J. Lee, A. Leminen, B. K. Lim, J. Lissowska, J. Lubinski, L. Lundvall, G. Lurie, L. F. Massuger, et al. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. *Nat Genet*, 45(4):362–70, 370e1–2, 2013.
- [96] S. E. Bojesen, K. A. Pooley, S. E. Johnatty, J. Beesley, K. Michailidou, J. P. Tyrer, S. L. Edwards, H. A. Pickett, H. C. Shen, C. E. Smart, K. M. Hillman, P. L. Mai, K. Lawrenson, M. D. Stutz, Y. Lu, R. Karevan, N. Woods, R. L. Johnston, J. D. French, X. Chen, M. Weischer, S. F. Nielsen, M. J. Maranian, M. Ghoussaini, S. Ahmed, C. Baynes, M. K. Bolla, Q. Wang, J. Dennis, L. McGuffog, D. Barrowdale, A. Lee, S. Healey, M. Lush, D. C. Tessier, D. Vincent, F. Bacot, Australian Cancer Study, Australian Ovarian Cancer Study Group, Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer, GENICA, Swedish Breast Cancer Study, Hereditary Breast and Ovarian Cancer Research Group Netherlands, Epidemiological study of BRCA Mutation Carriers, Genetic Modifiers of Cancer Risk in BRCA Mutation Carriers Study Collaborators, I. Vergote, S. Lambrechts, E. Despierre, H. A. Risch, A. Gonzalez-Neira, M. A. Rossing, G. Pita, J. A. Doherty, N. Alvarez, M. C. Larson, B. L. Fridley, N. Schoof, J. Chang-Claude, M. S. Cicek, J. Peto, K. R. Kalli, A. Brooks, S. M. Armasu, M. K. Schmidt, L. M. Braaf, B. Winterhoff, H. Nevanlinna, G. E. Konecny, D. Lambrechts, L. Rogmann, P. Guenel, A. Teoman, R. L. Milne, J. J. Garcia, A. Cox, V. Shridhar, B. Burwinkel, F. Marme, R. Hein, E. J. Sawyer, C. A. Haiman, S. Wang-Gohrke, I. L. Andrulis, K. B. Moysich, J. L. Hopper, K. Odunsi, A. Lindblom, G. G. Giles, H. Brenner, J. Simard, G. Lurie, P. A. Fasching, M. E. Carney, P. Radice, L. R. Wilkens, A. Swerdlow, M. T. Goodman, H. Brauch, M. Garcia-Closas, P. Hillemanns, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat Genet*, 45(4):371–84, 384e1–2, 2013.
- [97] J. Permuth-Wey, K. Lawrenson, H. C. Shen, A. Velkova, J. P. Tyrer, Z. Chen, H. Y. Lin, Y. A. Chen, Y. Y. Tsai, X. Qu, S. J. Ramus, R. Karevan, J. Lee, N. Lee, M. C. Larson, K. K. Aben, H. Anton-Culver, N. Antonenkova, A. C. Antoniou, S. M. Armasu, Australian Cancer Study, Australian Ovarian Cancer Study, F. Bacot, L. Baglietto, E. V. Bandera, J. Barnholtz-Sloan, M. W. Beckmann, M. J. Birrer, G. Bloom, N. Bogdanova, L. A. Brinton, A. Brooks-Wilson, R. Brown, R. Butzow, Q. Cai, I. Campbell, J. Chang-Claude, S. Chanock, G. Chenevix-Trench, J. Q. Cheng, M. S. Cicek, G. A. Coetzee, Consortium of Investigators of Modifiers of BRCA, L. S. Cook, F. J. Couch, D. W. Cramer, J. M. Cunningham, A. Dansonka-Mieszkowska, E. Despierre, J. A. Doherty, T. Dork, A. du Bois, M. Durst, D. F. Easton, D. Eccles, R. Edwards, A. B. Ekici,

REFERENCES

- P. A. Fasching, D. A. Fenstermacher, J. M. Flanagan, M. Garcia-Closas, A. Gentry-Maharaj, G. G. Giles, R. M. Glasspool, J. Gonzalez-Bosquet, M. T. Goodman, M. Gore, B. Gorski, J. Gronwald, P. Hall, M. K. Halle, P. Harter, F. Heitz, P. Hillemanns, M. Hoatlin, C. K. Hogdall, E. Hogdall, S. Hosono, A. Jakubowska, A. Jensen, H. Jim, K. R. Kalli, B. Y. Karlan, S. B. Kaye, L. E. Kelemen, L. A. Kiemeny, F. Kikkawa, G. E. Konecny, C. Krakstad, S. K. Kjaer, J. Kupryjanczyk, D. Lambrechts, S. Lambrechts, J. M. Lancaster, N. D. Le, A. Leminen, D. A. Levine, D. Liang, B. K. Lim, J. Lin, et al. Identification and molecular characterization of a new ovarian cancer susceptibility locus at 17q21.31. *Nat Commun*, 4:1627, 2013.
- [98] H. Shen, B. L. Fridley, H. Song, K. Lawrenson, J. M. Cunningham, S. J. Ramus, M. S. Cicek, J. Tyrer, D. Stram, M. C. Larson, M. Kobel, PRACTICAL Consortium, A. Ziogas, W. Zheng, H. P. Yang, A. H. Wu, E. L. Wozniak, Y. L. Woo, B. Winterhoff, E. Wik, A. S. Whittemore, N. Wentzensen, R. P. Weber, A. F. Vitonis, D. Vincent, R. A. Vierkant, I. Vergote, D. Van Den Berg, A. M. Van Altena, S. S. Tworoger, P. J. Thompson, D. C. Tessier, K. L. Terry, S. H. Teo, C. Templeman, D. O. Stram, M. C. Southey, W. Sieh, N. Siddiqui, Y. B. Shvetsov, X. O. Shu, V. Shridhar, S. Wang-Gohrke, G. Severi, I. Schwaab, H. B. Salvesen, I. K. Rzepecka, I. B. Runnebaum, M. A. Rossing, L. Rodriguez-Rodriguez, H. A. Risch, S. P. Renner, E. M. Poole, M. C. Pike, C. M. Phelan, L. M. Pelttari, T. Pejovic, J. Paul, I. Orlow, S. Z. Omar, S. H. Olson, K. Odunsi, S. Nickels, H. Nevanlinna, R. B. Ness, S. A. Narod, T. Nakanishi, K. B. Moysich, A. N. Monteiro, J. Moes-Sosnowska, F. Modugno, U. Menon, J. R. McLaughlin, V. McGuire, K. Matsuo, N. A. Adenani, L. F. Massuger, G. Lurie, L. Lundvall, J. Lubinski, J. Lissowska, D. A. Levine, A. Leminen, A. W. Lee, N. D. Le, S. Lambrechts, D. Lambrechts, J. Kupryjanczyk, C. Krakstad, G. E. Konecny, S. K. Kjaer, L. A. Kiemeny, L. E. Kelemen, G. L. Keeney, B. Y. Karlan, R. Karevan, K. R. Kalli, H. Kajiyama, B. T. Ji, A. Jensen, et al. Epigenetic analysis leads to identification of HNF1B as a subtype-specific susceptibility gene for ovarian cancer. *Nat Commun*, 4:1628, 2013.
- [99] K. Chen, H. Ma, L. Li, R. Zang, C. Wang, F. Song, T. Shi, D. Yu, M. Yang, W. Xue, J. Dai, S. Li, H. Zheng, C. Wu, Y. Zhang, X. Wu, D. Li, F. Xue, H. Li, Z. Jiang, J. Liu, Y. Liu, P. Li, W. Tan, J. Han, J. Jie, Q. Hao, Z. Hu, D. Lin, D. Ma, W. Jia, H. Shen, and Q. Wei. Genome-wide association study identifies new susceptibility loci for epithelial ovarian cancer in Han Chinese women. *Nat Commun*, 5:4682, 2014.
- [100] L. E. Kelemen, K. Lawrenson, J. Tyrer, Q. Li, J. M. Lee, J. H. Seo, C. M. Phelan, J. Beesley, X. Chen, T. J. Spindler, K. K. Aben, H. Anton-Culver, N. Antonenkova, Australian Cancer Study, Australian Ovarian Cancer Study Group, and Ovarian Cancer Association Consortium. Genome-wide significant risk associations for mucinous ovarian carcinoma. *Nat Genet*, 47(8):888–97, 2015.
- [101] M. Earp, S. J. Winham, N. Larson, J. B. Permuth, H. Sicotte, J. Chien, H. Anton-Culver, E. V. Bandera, A. Berchuck, L. S. Cook, D. Cramer, J. A. Doherty, M. T. Goodman, D. A. Levine, A. N. Monteiro, R. B. Ness, C. L. Pearce, M. A. Rossing, S. S. Tworoger, N. Wentzensen, M. Bisogna, L. Brinton, A. Brooks-Wilson, M. E. Carney, J. M. Cunningham, R. P. Edwards, Z. C. Fogarty, E. S. Iversen, P. Kraft, M. C. Larson, N. D. Le, H. Y. Lin, J. Lissowska, F. Modugno, K. B. Moysich, S. H. Olson, M. C. Pike, E. M. Poole, D. N. Rider, K. L. Terry, P. J. Thompson, D. van den Berg, R. A. Vierkant, A. F. Vitonis, L. R. Wilkens, A. H. Wu, H. P. Yang, A. Ziogas, C. M. Phelan, J. M. Schildkraut, Y. A. Chen, T. A. Sellers, B. L. Fridley, and E. L. Goode. A targeted genetic association study of epithelial ovarian cancer susceptibility. *Oncotarget*, 7(7):7381–9, 2016.
- [102] J. B. Permuth, A. Pirie, Y. A. Chen, H. Y. Lin, B. M. Reid, Z. Chen, A. Monteiro, J. Dennis, G. Mendoza-Fandino, Australian Ovarian Cancer Study Group, Australian Cancer Study, H. Anton-Culver, E. V. Bandera, M. Bisogna, L. Brinton, A. Brooks-Wilson, M. E. Carney, G. Chenevix-Trench, L. S. Cook, D. W. Cramer, J. M. Cunningham, C. Cybulski, A. A. D’Aloisio, J. A. Doherty, M. Earp, R. P. Edwards, B. L. Fridley, S. A. Gayther, A. Gentry-Maharaj, M. T. Goodman, J. Gronwald, E. Hogdall, E. S. Iversen, A. Jakubowska, A. Jensen, B. Y. Karlan, L. E. Kelemen, S. K. Kjaer, P. Kraft, N. D. Le, D. A. Levine, J. Lissowska, J. Lubinski, K. Matsuo,

- U. Menon, R. Modugno, K. B. Moysich, T. Nakanishi, R. B. Ness, S. Olson, I. Orlow, C. L. Pearce, T. Pejovic, E. M. Poole, S. J. Ramus, M. A. Rossing, D. P. Sandler, X. O. Shu, H. Song, J. A. Taylor, S. H. Teo, K. L. Terry, P. J. Thompson, S. S. Tworoger, P. M. Webb, N. Wentzensen, L. R. Wilkens, S. Winham, Y. L. Woo, A. H. Wu, H. Yang, W. Zheng, A. Ziogas, C. M. Phelan, J. M. Schildkraut, A. Berchuck, E. L. Goode, P. D. Pharoah, T. A. Sellers, and Ovarian Cancer Association Consortium. Exome genotyping arrays to identify rare and low frequency variants associated with epithelial ovarian cancer risk. *Hum Mol Genet*, 2016.
- [103] V. Beral, K. Gaitskell, C. Hermon, K. Moser, G. Reeves, and R. Peto. Ovarian cancer and smoking: individual participant meta-analysis including 28,114 women with ovarian cancer from 51 epidemiological studies. *Lancet Oncol*, 13(9):946–56, 2012.
- [104] K. Pirie, R. Peto, G. K. Reeves, J. Green, and V. Beral. The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. *Lancet*, 381(9861):133–41, 2013.
- [105] K. E. Bradbury, A. Balkwill, E. A. Spencer, A. W. Roddam, G. K. Reeves, J. Green, T. J. Key, V. Beral, K. Pirie, and Million Women Study Collaborators. Organic food consumption and the incidence of cancer in a large prospective study of women in the United Kingdom. *Br J Cancer*, 110(9):2321–6, 2014.
- [106] T. O. Yang, G. K. Reeves, J. Green, V. Beral, B. J. Cairns, and Million Women Study Collaborators. Birth weight and adult cancer incidence: large prospective study and meta-analysis. *Ann Oncol*, 25(9):1836–43, 2014.
- [107] H. O. Adami, C. C. Hsieh, M. Lambe, D. Trichopoulos, D. Leon, I. Persson, A. Ekblom, and P. O. Janson. Parity, age at first childbirth, and risk of ovarian cancer. *Lancet*, 344(8932):1250–4, 1994.
- [108] G. Albrechtsen, I. Heuch, and G. Kvale. Reproductive factors and incidence of epithelial ovarian cancer: a Norwegian prospective study. *Cancer Causes Control*, 7(4):421–7, 1996.
- [109] M. Booth, V. Beral, and P. Smith. Risk factors for ovarian cancer: a case-control study. *Br J Cancer*, 60(4):592–8, 1989.
- [110] M. G. Braem, N. C. Onland-Moret, P. A. van den Brandt, R. A. Goldbohm, P. H. Peeters, R. F. Kruitwagen, and L. J. Schouten. Reproductive and hormonal factors in association with ovarian cancer in the Netherlands cohort study. *Am J Epidemiol*, 172(10):1181–9, 2010.
- [111] M. A. Gates, B. A. Rosner, J. L. Hecht, and S. S. Tworoger. Risk factors for epithelial ovarian cancer by histologic subtype. *Am J Epidemiol*, 171(1):45–53, 2010.
- [112] A. W. Kurian, R. R. Balise, V. McGuire, and A. S. Whittemore. Histologic types of epithelial ovarian cancer: have they different risk factors? *Gynecol Oncol*, 96(2):520–30, 2005.
- [113] F. Modugno, R. B. Ness, and J. E. Wheeler. Reproductive risk factors for epithelial ovarian cancer according to histologic type and invasiveness. *Ann Epidemiol*, 11(8):568–74, 2001.
- [114] P. G. Moorman, R. T. Palmieri, L. Akushevich, A. Berchuck, and J. M. Schildkraut. Ovarian cancer risk factors in African-American and white women. *Am J Epidemiol*, 170(5):598–606, 2009.
- [115] E. Negri, S. Franceschi, A. Tzonou, M. Booth, C. La Vecchia, F. Parazzini, V. Beral, P. Boyle, and D. Trichopoulos. Pooled analysis of 3 European case-control studies: I. Reproductive factors and risk of epithelial ovarian cancer. *Int J Cancer*, 49(1):50–6, 1991.
- [116] M. C. Pike, C. L. Pearce, R. Peters, W. Cozen, P. Wan, and A. H. Wu. Hormonal factors and the risk of invasive ovarian cancer: a population-based case-control study. *Fertil Steril*, 82(1):186–95, 2004.

REFERENCES

- [117] D. Purdie, A. Green, C. Bain, V. Siskind, B. Ward, N. Hacker, M. Quinn, G. Wright, P. Russell, and B. Susil. Reproductive and other factors and risk of epithelial ovarian cancer: an Australian case-control study. Survey of Women's Health Study Group. *Int J Cancer*, 62(6):678–84, 1995.
- [118] T. Riman, P. W. Dickman, S. Nilsson, N. Correia, H. Nordlinder, C. M. Magnusson, and I. R. Persson. Risk factors for invasive epithelial ovarian cancer: results from a Swedish case-control study. *Am J Epidemiol*, 156(4):363–73, 2002.
- [119] H. A. Risch, L. D. Marrett, M. Jain, and G. R. Howe. Differences in risk factors for epithelial ovarian cancer by histologic type. Results of a case-control study. *Am J Epidemiol*, 144(4):363–72, 1996.
- [120] L. Titus-Ernstoff, K. Perez, D. W. Cramer, B. L. Harlow, J. A. Baron, and E. R. Greenberg. Menstrual and reproductive factors in relation to ovarian cancer risk. *Br J Cancer*, 84(5):714–21, 2001.
- [121] K. H. Tung, M. T. Goodman, A. H. Wu, K. McDuffie, L. R. Wilkens, L. N. Kolonel, A. M. Nomura, K. Y. Terada, M. E. Carney, and L. H. Sobin. Reproductive factors and epithelial ovarian cancer risk by histologic type: a multiethnic case-control study. *Am J Epidemiol*, 158(7):629–38, 2003.
- [122] H. P. Yang, B. Trabert, M. A. Murphy, M. E. Sherman, J. N. Sampson, L. A. Brinton, P. Hartge, A. Hollenbeck, Y. Park, and N. Wentzensen. Ovarian cancer risk factors by histologic subtypes in the NIH-AARP Diet and Health Study. *Int J Cancer*, 131(4):938–48, 2012.
- [123] L. P. Feng, H. L. Chen, and M. Y. Shen. Breastfeeding and the risk of ovarian cancer: a meta-analysis. *J Midwifery Womens Health*, 59(4):428–37, 2014.
- [124] D. P. Li, C. Du, Z. M. Zhang, G. X. Li, Z. F. Yu, X. Wang, P. F. Li, C. Cheng, Y. P. Liu, and Y. S. Zhao. Breastfeeding and ovarian cancer risk: a systematic review and meta-analysis of 40 epidemiological studies. *Asian Pac J Cancer Prev*, 15(12):4829–37, 2014.
- [125] N. N. Luan, Q. J. Wu, T. T. Gong, E. Vogtmann, Y. L. Wang, and B. Lin. Breastfeeding and ovarian cancer risk: a meta-analysis of epidemiologic studies. *Am J Clin Nutr*, 98(4):1020–31, 2013.
- [126] World Cancer Research Fund / American Institute for Cancer Research. Food, nutrition, physical activity, and the prevention of ovarian cancer. *Continuous Update Project Report*, Available at: http://www.dietandcancerreport.org/cup/cup_resources.php, 2014.
- [127] D. Cibula, M. Widschwendter, O. Majek, and L. Dusek. Tubal ligation and the risk of ovarian cancer: review and meta-analysis. *Hum Reprod Update*, 17(1):55–67, 2011.
- [128] C. L. Pearce, C. Templeman, M. A. Rossing, A. Lee, A. M. Near, P. M. Webb, C. M. Nagle, J. A. Doherty, K. L. Cushing-Haugen, K. G. Wicklund, J. Chang-Claude, R. Hein, G. Lurie, L. R. Wilkens, M. E. Carney, M. T. Goodman, K. Moysich, S. K. Kjaer, E. Hogdall, A. Jensen, E. L. Goode, B. L. Fridley, M. C. Larson, J. M. Schildkraut, R. T. Palmieri, D. W. Cramer, K. L. Terry, A. F. Vitonis, L. J. Titus, A. Ziogas, W. Brewster, H. Anton-Culver, A. Gentry-Maharaj, S. J. Ramus, A. R. Anderson, D. Brueggmann, P. A. Fasching, S. A. Gayther, D. G. Huntsman, U. Menon, R. B. Ness, M. C. Pike, H. Risch, A. H. Wu, and A. Berchuck. Association between endometriosis and risk of histological subtypes of ovarian cancer: a pooled analysis of case-control studies. *Lancet Oncol*, 13(4):385–94, 2012.
- [129] S. J. Jordan, C. M. Nagle, M. D. Coory, D. Maresco, M. M. Protani, N. A. Pandeya, K. D. Balasubramaniam, and P. M. Webb. Has the association between hysterectomy and ovarian cancer changed over time? A systematic review and meta-analysis. *Eur J Cancer*, 49(17):3638–47, 2013.

-
- [130] I. Rizzuto, R. F. Behrens, and L. A. Smith. Risk of ovarian cancer in women treated with ovarian stimulating drugs for infertility. *Cochrane Database Syst Rev*, (8):CD008215, 2013.
- [131] T. E. Crane, B. R. Khulpateea, D. S. Alberts, K. Basen-Engquist, and C. A. Thomson. Dietary intake and ovarian cancer risk: a systematic review. *Cancer Epidemiol Biomarkers Prev*, 23(2): 255–73, 2014.
- [132] M. A. Merritt, I. Tzoulaki, P. A. van den Brandt, L. J. Schouten, K. K. Tsilidis, E. Weiderpass, C. J. Patel, A. Tjonneland, L. Hansen, K. Overvad, M. His, L. Dartois, M. C. Boutron-Ruault, R. T. Fortner, R. Kaaks, K. Aleksandrova, H. Boeing, A. Trichopoulou, P. Lagiou, C. Bamia, D. Palli, V. Krogh, R. Tumino, F. Ricceri, A. Mattiello, H. B. Bueno-de Mesquita, N. C. Onland-Moret, P. H. Peeters, G. Skeie, M. Jareid, J. R. Quiros, M. Obon-Santacana, M. J. Sanchez, S. Chamosa, J. M. Huerta, A. Barricarte, J. A. Dias, E. Sonestedt, A. Idahl, E. Lundin, N. J. Wareham, K. T. Khaw, R. C. Travis, P. Ferrari, E. Riboli, and M. J. Gunter. Nutrient-wide association study of 57 foods/nutrients and epithelial ovarian cancer in the European Prospective Investigation into Cancer and Nutrition study and the Netherlands Cohort Study. *Am J Clin Nutr*, 103(1):161–7, 2016.
- [133] K. L. Terry, S. Karageorgi, Y. B. Shvetsov, M. A. Merritt, G. Lurie, P. J. Thompson, M. E. Carney, R. P. Weber, L. Akushevich, W. H. Lo-Ciganic, K. Cushing-Haugen, W. Sieh, K. Moysich, J. A. Doherty, C. M. Nagle, A. Berchuck, C. L. Pearce, M. Pike, R. B. Ness, P. M. Webb, Australian Cancer Study, Australian Ovarian Cancer Study Group, M. A. Rossing, J. Schildkraut, H. Risch, M. T. Goodman, and Ovarian Cancer Association Consortium. Genital powder use and risk of ovarian cancer: a pooled analysis of 8,525 cases and 9,859 controls. *Cancer Prev Res (Phila)*, 6(8):811–21, 2013.
- [134] D. M. Gertig, D. J. Hunter, D. W. Cramer, G. A. Colditz, F. E. Speizer, W. C. Willett, and S. E. Hankinson. Prospective study of talc use and ovarian cancer. *J Natl Cancer Inst*, 92(3):249–52, 2000.
- [135] S. C. Houghton, K. W. Reeves, S. E. Hankinson, L. Crawford, D. Lane, J. Wactawski-Wende, C. A. Thomson, J. K. Ockene, and S. R. Sturgeon. Perineal powder use and risk of ovarian cancer. *J Natl Cancer Inst*, 106(9), 2014.
- [136] T. T. Gong, Q. J. Wu, E. Vogtmann, B. Lin, and Y. L. Wang. Age at menarche and risk of ovarian cancer: a meta-analysis of epidemiological studies. *Int J Cancer*, 132(12):2894–900, 2013.
- [137] T. M. Brasky, J. Liu, E. White, U. Peters, J. D. Potter, R. B. Walter, C. S. Baik, D. S. Lane, J. E. Manson, M. Z. Vitolins, M. A. Allison, J. Y. Tang, and J. Wactawski-Wende. Non-steroidal anti-inflammatory drugs and cancer risk in women: results from the Women’s Health Initiative. *Int J Cancer*, 135(8):1869–83, 2014.
- [138] B. Trabert, R. B. Ness, W. H. Lo-Ciganic, M. A. Murphy, E. L. Goode, E. M. Poole, L. A. Brinton, P. M. Webb, C. M. Nagle, S. J. Jordan, Australian Ovarian Cancer Study Group, Australian Cancer Study, H. A. Risch, M. A. Rossing, J. A. Doherty, M. T. Goodman, G. Lurie, S. K. Kjaer, E. Hogdall, A. Jensen, D. W. Cramer, K. L. Terry, A. Vitonis, E. V. Bandera, S. Olson, M. G. King, U. Chandran, H. Anton-Culver, A. Ziogas, U. Menon, S. A. Gayther, S. J. Ramus, A. Gentry-Maharaj, A. H. Wu, C. L. Pearce, M. C. Pike, A. Berchuck, J. M. Schildkraut, N. Wentzensen, and Ovarian Cancer Association Consortium. Aspirin, nonaspirin nonsteroidal anti-inflammatory drug, and acetaminophen use and risk of invasive epithelial ovarian cancer: a pooled analysis in the Ovarian Cancer Association Consortium. *J Natl Cancer Inst*, 106(2):djt431, 2014.
- [139] M. A. Murphy, B. Trabert, H. P. Yang, Y. Park, L. A. Brinton, P. Hartge, M. E. Sherman, A. Hollenbeck, and N. Wentzensen. Non-steroidal anti-inflammatory drug use and ovarian cancer risk: findings from the NIH-AARP Diet and Health Study and systematic review. *Cancer Causes Control*, 23(11):1839–52, 2012.

REFERENCES

- [140] M. F. Fathalla. Incessant ovulation—a factor in ovarian neoplasia? *Lancet*, 2(7716):163, 1971.
- [141] M. F. Fathalla. Incessant ovulation and ovarian cancer - a hypothesis re-visited. *Facts Views Vis Obgyn*, 5(4):292–7, 2013.
- [142] A. Mantovani, P. Allavena, A. Sica, and F. Balkwill. Cancer-related inflammation. *Nature*, 454(7203):436–44, 2008.
- [143] K. A. Lang Kuhs, A. Hildesheim, B. Trabert, T. J. Kemp, M. P. Purdue, N. Wentzensen, H. A. Katki, L. A. Pinto, E. Loftfield, M. Safaeian, A. K. Chaturvedi, and M. S. Shiels. Association between regular aspirin use and circulating markers of inflammation: A study within the Prostate, Lung, Colorectal, and Ovarian cancer screening trial. *Cancer Epidemiol Biomarkers Prev*, 24(5):825–32, 2015.
- [144] R. B. Ness and C. Cottreau. Possible role of ovarian epithelial inflammation in ovarian cancer. *J Natl Cancer Inst*, 91(17):1459–67, 1999.
- [145] R. B. Ness, J. A. Grisso, C. Cottreau, J. Klapper, R. Vergona, J. E. Wheeler, M. Morgan, and J. J. Schlesselman. Factors related to inflammation of the ovarian epithelium and risk of ovarian cancer. *Epidemiology*, 11(2):111–7, 2000.
- [146] M. A. Merritt, A. C. Green, C. M. Nagle, P. M. Webb, Australian Cancer Study, and Australian Ovarian Cancer Study Group. Talcum powder, chronic pelvic inflammation and NSAIDs in relation to risk of epithelial ovarian cancer. *Int J Cancer*, 122(1):170–6, 2008.
- [147] The Million Women Study Collaborative Group. The Million Women Study: design and characteristics of the study population. *Breast Cancer Res*, 1(1):73–80, 1999.
- [148] Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet*, 362(9382):419–27, 2003.
- [149] Hospital Episode Statistics. <http://www.hscic.gov.uk/hes>.
- [150] S. Kendrick and J. Clarke. The Scottish record linkage system. *Health Bull (Edinb)*, 51(2):72–9, 1993.
- [151] World Health Organization. *International statistical classification of diseases and related health problems*. World Health Organization, Geneva, 10th revision (ICD-10) edition, 1992.
- [152] World Health Organization. *Manual of the international statistical classification of diseases, injuries, and causes of death*. World Health Organization, Geneva, 9th revision (ICD-9) edition, 1977.
- [153] Office of Population Censuses and Surveys. *Tabular list of the classification of surgical operations and procedures, 4th revision*. HMSO, London, 1990.
- [154] C. Percy, V. Van Holten, and C. S. Muir (Eds.). *International Classification of Diseases for Oncology*. World Health Organization, Geneva, 2nd edition, 1990.
- [155] A. Fritz, C. Percy, A. Jack, K. Shanmugaratnam, L. Sobin, D.M. Parkin, and S. Whelan (Eds.). *International Classification of Diseases for Oncology*. World Health Organization, Geneva, 3rd edition, 2000.
- [156] K. Gaitskell, J. Green, K. Pirie, G. Reeves, V. Beral, and on behalf of the Million Women Study Collaborators. Tubal ligation and ovarian cancer risk in a large cohort: Substantial variation by histological type. *Int J Cancer*, 138(5):1076–84, 2016.

-
- [157] P. Townsend, P. Phillimore, and A. Beattie. *Health and Deprivation: Inequality and the North*. Croom Helm, London, 1988.
- [158] X. Xue, M. Y. Kim, M. M. Gaudet, Y. Park, M. Heo, A. R. Hollenbeck, H. D. Strickler, and M. J. Gunter. A comparison of the polytomous logistic regression and joint Cox proportional hazards models for evaluating multiple disease subtypes in prospective cohort studies. *Cancer Epidemiol Biomarkers Prev*, 22(2):275–85, 2013.
- [159] D. F. Easton, J. Peto, and A. G. Babiker. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med*, 10(7):1025–35, 1991.
- [160] M. Plummer. Improved estimates of floating absolute risk. *Stat Med*, 23(1):93–104, 2004.
- [161] StataCorp. Stata statistical software: Release 14, 2015.
- [162] R Development Core Team. R: A language and environment for statistical computing, 2015.
- [163] Matthew Arnold. Jasper: Jasper makes plots. R package version 2-192., 2015.
- [164] D Lader. *Opinions Survey report No.41: Contraception and Sexual Health, 2008/09*. Office for National Statistics, London, 2009.
- [165] L Rainford and H Meltzer. *Contraception and Sexual Health, 1998*. Office for National Statistics, London, 2000.
- [166] F Dawe and H Meltzer. *Contraception and Sexual Health, 2002*. Office for National Statistics, London, 2003.
- [167] M. S. Rice, M. A. Murphy, and S. S. Tworoger. Tubal ligation, hysterectomy and ovarian cancer: A meta-analysis. *J Ovarian Res*, 5(1):13, 2012.
- [168] J. H. Shepherd. Revised FIGO staging for gynaecological cancer. *Br J Obstet Gynaecol*, 96(8): 889–92, 1989.
- [169] F. A. Tavassoli and P. Devilee. *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Breast and Female Genital Organs*. IARC Press, Lyon, 2003.
- [170] S. G. Silverberg. Histopathologic grading of ovarian carcinoma: a review and proposal. *Int J Gynecol Pathol*, 19(1):7–15, 2000.
- [171] D. C. Bodurka, M. T. Deavers, C. Tian, C. C. Sun, A. Malpica, R. L. Coleman, K. H. Lu, A. K. Sood, M. J. Birrer, R. Ozols, R. Baergen, R. E. Emerson, M. Steinhoff, B. Behmaram, G. Rasty, and D. M. Gershenson. Reclassification of serous ovarian carcinoma by a 2-tier system: a Gynecologic Oncology Group Study. *Cancer*, 118(12):3087–94, 2012.
- [172] J. Prat and for the FIGO Committee on Gynecologic Oncology. FIGO guidelines: Staging classification for cancer of the ovary, fallopian tube, and peritoneum. *Int J Gynaecol Obstet*, 124(1): 1–5, 2014.
- [173] N. Wilkinson, W. G. McCluggage, and on behalf of the Royal College of Pathologists’ Working Group on Cancer Services. Datasets for the histopathological reporting of neoplasms of the ovaries and fallopian tubes and primary carcinomas of the peritoneum (3rd edition). *Royal College of Pathologists: Standards and datasets for reporting cancers*, 2010.
- [174] D. Cibula, M. Widschwendter, M. Zikan, and L. Dusek. Underlying mechanisms of ovarian cancer risk reduction after tubal ligation. *Acta Obstet Gynecol Scand*, 90(6):559–63, 2011.

REFERENCES

- [175] C. Madsen, L. Baandrup, C. Dehlendorff, and S. K. Kjaer. Tubal ligation and salpingectomy and the risk of epithelial ovarian cancer and borderline ovarian tumors: a nationwide case-control study. *Acta Obstet Gynecol Scand*, 94(1):86–94, 2015.
- [176] I. T. Gram, A. Lukanova, I. Brill, T. Braaten, E. Lund, E. Lundin, K. Overvad, A. Tjonneland, F. Clavel-Chapelon, N. Chabbert-Buffet, C. Bamia, A. Trichopoulou, D. Zylis, G. Masala, F. Berrino, R. Galasso, R. Tumino, C. Sacerdote, O. Gavrilyuk, S. Kristiansen, L. Rodriguez, C. Bonet, J. M. Huerta, A. Barricarte, M. J. Sanchez, M. Dorronsoro, K. Jirstrom, M. Almquist, A. Idahl, H. B. Bueno-de Mesquita, M. Braem, C. Onland-Moret, K. K. Tsilidis, N. E. Allen, V. Fedirko, E. Riboli, and R. Kaaks. Cigarette smoking and risk of histological subtypes of epithelial ovarian cancer in the EPIC cohort study. *Int J Cancer*, 130(9):2204–10, 2012.
- [177] M. Kobel, S. E. Kalloger, D. G. Huntsman, J. L. Santos, K. D. Swenerton, J. D. Seidman, and C. B. Gilks. Differences in tumor type in low-stage versus high-stage ovarian carcinomas. *Int J Gynecol Pathol*, 29(3):203–11, 2010.
- [178] J. D. Seidman, I. Horkayne-Szakaly, M. Haiba, C. R. Boice, R. J. Kurman, and B. M. Ronnett. The histologic type and stage distribution of ovarian carcinomas of surface epithelial origin. *Int J Gynecol Pathol*, 23(1):41–4, 2004.
- [179] K. Gaitskell, K. Coffey, J. Green, K. Pirie, G. K. Reeves, A. A. Ahmed, I. Barnes, V. Beral, and on behalf of the Million Women Study Collaborators. Tubal ligation and incidence of 26 site-specific cancers in the Million Women Study. *Br J Cancer*, 114(9):1033–7, 2016.
- [180] K. L. Irwin, N. C. Lee, H. B. Peterson, G. L. Rubin, P. A. Wingo, and M. G. Mandel. Hysterectomy, tubal sterilization, and the risk of breast cancer. *Am J Epidemiol*, 127(6):1192–201, 1988.
- [181] N. Kreiger, M. Sloan, M. Cotterchio, and V. Kirsh. The risk of breast cancer following reproductive surgery. *Eur J Cancer*, 35(1):97–101, 1999.
- [182] S. K. Kjaer, L. Møller, L. A. Brinton, C. Johansen, G. Gridley, and J. H. Olsen. Tubal sterilization and risk of ovarian, endometrial and cervical cancer. A Danish population-based follow-up study of more than 65 000 sterilized women. *Int J Epidemiol*, 33(3):596–602, 2004.
- [183] C. A. Mathews, J. A. Stoner, N. Wentzensen, K. M. Moxley, M. E. Tenney, E. R. Tuller, T. Myers, L. M. Landrum, G. Lanneau, R. E. Zuna, M. A. Gold, S. S. Wang, and J. L. Walker. Tubal ligation frequency in Oklahoma women with cervical cancer. *Gynecol Oncol*, 127(2):278–82, 2012.
- [184] T. Dorjgochoo, X. O. Shu, H. L. Li, H. Z. Qian, G. Yang, H. Cai, Y. T. Gao, and W. Zheng. Use of oral contraceptives, intrauterine devices and tubal sterilization and cancer risk in a large prospective study, from 1996 to 2006. *Int J Cancer*, 124(10):2442–9, 2009.
- [185] D. J. Grant, P. G. Moorman, L. Akushevich, R. T. Palmieri, R. C. Bentley, and J. M. Schildkraut. Primary peritoneal and ovarian cancers: an epidemiological comparative analysis. *Cancer Causes Control*, 21(7):991–8, 2010.
- [186] A. Green, D. Purdie, C. Bain, V. Siskind, P. Russell, M. Quinn, and B. Ward. Tubal sterilisation, hysterectomy and decreased risk of ovarian cancer. Survey of Women’s Health Study Group. *Int J Cancer*, 71(6):948–51, 1997.
- [187] A. Riska, R. Sund, E. Pukkala, M. Gissler, and A. Leminen. Parity, tubal sterilization, hysterectomy and risk of primary fallopian tube carcinoma in Finland, 1975-2004. *Int J Cancer*, 120(6):1351–4, 2007.

-
- [188] D. Vicus, A. Finch, B. Rosen, I. Fan, L. Bradley, I. Cass, P. Sun, B. Karlan, J. McLaughlin, S. A. Narod, and Hereditary Ovarian Cancer Clinical Study Group. Risk factors for carcinoma of the fallopian tube in women with and without a germline BRCA mutation. *Gynecol Oncol*, 118(2): 155–9, 2010.
- [189] N. N. Nik, R. Vang, Ie M. Shih, and R. J. Kurman. Origin and pathogenesis of pelvic (ovarian, tubal, and primary peritoneal) serous carcinoma. *Annu Rev Pathol*, 9:27–45, 2014.
- [190] R. Perets and R. Drapkin. It’s totally tubular...riding the new wave of ovarian cancer research. *Cancer Res*, 76(1):10–7, 2016.
- [191] H. Falconer, L. Yin, H. Gronberg, and D. Altman. Ovarian cancer risk after salpingectomy: a nationwide population-based study. *J Natl Cancer Inst*, 107(2), 2015.
- [192] C. R. Lessard-Anderson, K. S. Handlogten, R. J. Molitor, S. C. Dowdy, W. A. Cliby, A. L. Weaver, J. S. Sauver, and J. N. Bakkum-Gamez. Effect of tubal sterilization technique on risk of serous epithelial ovarian and primary peritoneal carcinoma. *Gynecol Oncol*, 135(3):423–7, 2014.
- [193] S. H. Yoon, S. N. Kim, S. H. Shim, S. B. Kang, and S. J. Lee. Bilateral salpingectomy can reduce the risk of ovarian cancer in the general population: A meta-analysis. *Eur J Cancer*, 55:38–46, 2016.
- [194] X. Castellsague, W. D. Thompson, and R. Dubrow. Tubal sterilization and the risk of endometrial cancer. *Int J Cancer*, 65(5):607–12, 1996.
- [195] J. V. Lacey, Jr., L. A. Brinton, R. Mortel, M. L. Berman, G. D. Wilbanks, L. B. Twiggs, and R. J. Barrett. Tubal sterilization and risk of cancer of the endometrium. *Gynecol Oncol*, 79(3):482–4, 2000.
- [196] H. Li and D. B. Thomas. Tubal ligation and risk of cervical cancer. The World Health Organization collaborative study of neoplasia and steroid contraceptives. *Contraception*, 61(5):323–8, 2000.
- [197] E. E. Calle, C. Rodriguez, K. A. Walker, P. A. Wingo, J. M. Petrelli, and M. J. Thun. Tubal sterilization and risk of breast cancer mortality in US women. *Cancer Causes Control*, 12(2): 127–35, 2001.
- [198] L. A. Brinton, M. D. Gammon, R. J. Coates, and R. N. Hoover. Tubal ligation and risk of breast cancer. *Br J Cancer*, 82(9):1600–4, 2000.
- [199] A. H. Eliassen, G. A. Colditz, B. Rosner, and S. E. Hankinson. Tubal sterilization in relation to breast cancer risk. *Int J Cancer*, 118(8):2026–30, 2006.
- [200] M. M. Gaudet, A. V. Patel, J. Sun, L. R. Teras, and S. M. Gapstur. Tubal sterilization and breast cancer incidence: results from the Cancer Prevention Study II nutrition cohort and meta-analysis. *Am J Epidemiol*, 177(6):492–9, 2013.
- [201] L. Iversen, P. C. Hannaford, and A. M. Elliott. Tubal sterilization, all-cause death, and cancer among women in the United Kingdom: evidence from the Royal College of General Practitioners’ Oral Contraception Study. *Am J Obstet Gynecol*, 196(5):447 e1–8, 2007.
- [202] H. B. Nichols, D. D. Baird, L. A. DeRoo, G. E. Kissling, and D. P. Sandler. Tubal ligation in relation to menopausal symptoms and breast cancer risk. *Br J Cancer*, 109(5):1291–5, 2013.
- [203] D. J. Press, J. Sullivan-Halley, G. Ursin, D. Deapen, J. A. McDonald, B. L. Strom, S. A. Norman, M. S. Simon, P. A. Marchbanks, S. G. Folger, J. M. Liff, R. T. Burkman, K. E. Malone, L. K. Weiss, R. Spirtas, and L. Bernstein. Breast cancer risk and ovariectomy, hysterectomy, and tubal sterilization in the women’s contraceptive and reproductive experiences study. *Am J Epidemiol*, 173(1):38–47, 2011.

REFERENCES

- [204] J. M. Walboomers, M. V. Jacobs, M. M. Manos, F. X. Bosch, J. A. Kummer, K. V. Shah, P. J. Snijders, J. Peto, C. J. Meijer, and N. Munoz. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol*, 189(1):12–19, 1999.
- [205] International Agency for Research on Cancer. *A review of human carcinogens. Part B: Biological agents: Human Papillomaviruses*, volume 100B of *IARC Monographs on the evaluation of carcinogenic risks to humans*. IARC Press, Lyon, France, 2009.
- [206] D. B. Cape and N. Kreiger. Gynaecological surgical procedures and risk of colorectal cancer in women. *Eur J Cancer Prev*, 8(6):495–500, 1999.
- [207] K. A. Rosenblatt, D. L. Gao, R. M. Ray, Z. C. Nelson, and D. B. Thomas. Contraceptive methods and induced abortions and their association with the risk of colon cancer in Shanghai, China. *Eur J Cancer*, 40(4):590–3, 2004.
- [208] M. Z. Braganza, A. B. de Gonzalez, S. J. Schonfeld, N. Wentzensen, A. V. Brenner, and C. M. Kitahara. Benign breast and gynecologic conditions, reproductive and hormonal factors, and risk of thyroid cancer. *Cancer Prev Res (Phila)*, 7(4):418–25, 2014.
- [209] A. U. Hakverdi, C. E. Taner, A. C. Erden, and O. Satıcı. Changes in ovarian function after tubal sterilization. *Adv Contracept*, 10(1):51–6, 1994.
- [210] E. Radwanska, S. K. Headley, and P. Dmowski. Evaluation of ovarian function after tubal sterilization. *J Reprod Med*, 27(7):376–84, 1982.
- [211] J. Cattanach. Oestrogen deficiency after tubal ligation. *Lancet*, 1(8433):847–9, 1985.
- [212] J. Prat. Pathology of cancers of the female genital tract. *Int J Gynaecol Obstet*, 131 Suppl 2: S132–45, 2015.
- [213] R. T. Fortner, J. Ose, M. A. Merritt, H. Schock, A. Tjonneland, L. Hansen, K. Overvad, L. Dossus, F. Clavel-Chapelon, L. Baglietto, H. Boeing, A. Trichopoulou, V. Benetou, P. Lagiou, C. Agnoli, A. Mattiello, G. Masala, R. Tumino, C. Sacerdote, H. B. Bueno-de Mesquita, N. C. Onland-Moret, P. H. Peeters, E. Weiderpass, I. Torhild Gram, E. J. Duell, N. Larranaga, E. Ardanaz, M. J. Sanchez, M. D. Chirlaque, J. Brandstedt, A. Idahl, E. Lundin, K. T. Khaw, N. Wareham, R. C. Travis, S. Rinaldi, I. Romieu, M. J. Gunter, E. Riboli, and R. Kaaks. Reproductive and hormone-related risk factors for epithelial ovarian cancer by histologic pathways, invasiveness and histologic subtypes: Results from the EPIC cohort. *Int J Cancer*, 137(5):1196–208, 2015.
- [214] F. Chiaffarino, C. Pelucchi, F. Parazzini, E. Negri, S. Franceschi, R. Talamini, E. Conti, M. Montella, and C. La Vecchia. Reproductive and hormonal factors and ovarian cancer. *Ann Oncol*, 12(3):337–41, 2001.
- [215] D. W. Cramer, G. B. Hutchison, W. R. Welch, R. E. Scully, and K. J. Ryan. Determinants of ovarian cancer risk. I. Reproductive experiences and family history. *J Natl Cancer Inst*, 71(4): 711–6, 1983.
- [216] C. M. Vachon, P. J. Mink, C. A. Janney, T. A. Sellers, J. R. Cerhan, L. Hartmann, and A. R. Folsom. Association of parity and ovarian cancer risk by family history of breast or ovarian cancer in a population-based study of postmenopausal women. *Epidemiology*, 13(1):66–71, 2002.
- [217] M. Bevier, J. Sundquist, and K. Hemminki. Does the time interval between first and last birth influence the risk of endometrial and ovarian cancer? *Eur J Cancer*, 47(4):586–91, 2011.
- [218] G. S. Cooper, J. M. Schildkraut, A. S. Whittemore, and P. A. Marchbanks. Pregnancy recency and risk of ovarian cancer. *Cancer Causes Control*, 10(5):397–402, 1999.

-
- [219] D. C. Whiteman, V. Siskind, D. M. Purdie, and A. C. Green. Timing of pregnancy and the risk of epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev*, 12(1):42–6, 2003.
 - [220] C. Y. Yang, H. W. Kuo, and H. F. Chiu. Age at first birth, parity, and risk of death from ovarian cancer in Taiwan: a country of low incidence of ovarian cancer. *Int J Gynecol Cancer*, 17(1):32–6, 2007.
 - [221] M. Hickey, K. Ballard, and C. Farquhar. Endometriosis. *BMJ*, 348:g1752, 2014.
 - [222] H. A. Risch. Hormonal etiology of epithelial ovarian cancer, with a hypothesis concerning the role of androgens and progesterone. *J Natl Cancer Inst*, 90(23):1774–86, 1998.
 - [223] T. E. Bartlett, K. Chindera, J. McDermott, C. E. Breeze, W. R. Cooke, A. Jones, D. Reisel, S. T. Karegodar, R. Arora, S. Beck, U. Menon, L. Dubeau, and M. Widschwendter. Epigenetic reprogramming of fallopian tube fimbriae in BRCA mutation carriers defines early ovarian cancer evolution. *Nat Commun*, 7:11620, 2016.
 - [224] K. Hellner, F. Miranda, D. Fotso Chedom, S. Herrero-Gonzalez, D. M. Hayden, R. Tearle, M. Artibani, M. KaramiNejadRanjbar, R. Williams, K. Gaitskell, S. Elorbany, R. Xu, A. Laios, P. Buiga, K. Ahmed, S. Dhar, R. Y. Zhang, L. Campo, K. A. Myers, M. Lozano, M. Ruiz-Miro, S. Gatiús, A. Mota, G. Moreno-Bueno, X. Matias-Guiu, J. Benitez, L. Witty, G. McVean, S. Leedham, I. Tomlinson, R. Drmanac, J. B. Cazier, R. Klein, K. Dunne, R. C. Bast, Jr., S. H. Kennedy, B. Hassan, S. Lise, M. J. Garcia, B. A. Peters, C. Yau, T. Sauka-Spengler, and A. A. Ahmed. Premalignant SOX2 overexpression in the fallopian tubes of ovarian cancer patients: Discovery and validation studies. *EBioMedicine*, 10:137–49, 2016.
 - [225] S. H. George, A. Milea, and P. A. Shaw. Proliferation in the normal FTE is a hallmark of the follicular phase, not BRCA mutation status. *Clin Cancer Res*, 18(22):6199–207, 2012.
 - [226] B. M. Norquist, R. L. Garcia, K. H. Allison, C. H. Jokinen, L. E. Kernochan, C. C. Pizzi, B. J. Barrow, B. A. Goff, and E. M. Swisher. The molecular pathogenesis of hereditary ovarian carcinoma: alterations in the tubal epithelium of women with BRCA1 and BRCA2 mutations. *Cancer*, 116(22):5261–71, 2010.
 - [227] A. Anton, S. Schott, G. Kaip, M. Rath, J. Heil, S. Aulmann, and H. P. Sinn. Ki-67 and p53 expression of the fallopian tube mucosa in breast cancer patients with hereditary risk. *Arch Gynecol Obstet*, 289(5):1079–85, 2014.
 - [228] A. E. Rizzardi, A. T. Johnson, R. I. Vogel, S. E. Pambuccian, J. Henriksen, A. P. Skubitz, G. J. Metzger, and S. C. Schmechel. Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. *Diagn Pathol*, 7:42, 2012.
 - [229] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. FIJI: an open-source platform for biological-image analysis. *Nat Methods*, 9(7):676–82, 2012.
 - [230] J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1(8476):307–10, 1986.
 - [231] C. Fotopoulou, M. Hall, S. Kehoe, R. Ganesan, D. Cruickshank, C. Hughes, S. Sundar, H. Gabra, and R. Naik. *BGCS Guidelines for Epithelial Ovarian / Fallopian tube / Primary Peritoneal Cancer: Recommendations for Practice (Draft Guidelines)*. British Gynaecological Cancer Society, URL: <https://bgcs.org.uk/news/ovarian-and-endometrial-cancer-guidelines.html> (Accessed 22/07/2016), 2016.

REFERENCES

- [232] I. J. G. Harley, J. Quinn, J. P. Beirne, and W. G. McCluggage. *Scientific Impact Paper No. 44: The distal fallopian tube as the origin of non-uterine pelvic high-grade serous carcinoma*. Scientific Impact Paper. Royal College of Obstetricians and Gynaecologists, London, 2014.
- [233] Society of Gynecologic Oncology. *SGO Clinical Practice Statement: Salpingectomy for Ovarian Cancer Prevention*. SGO, Chicago, 2013.
- [234] M. H. Greene, P. L. Mai, and P. E. Schwartz. Does bilateral salpingectomy with ovarian retention warrant consideration as a temporary bridge to risk-reducing bilateral oophorectomy in BRCA1/2 mutation carriers? *Am J Obstet Gynecol*, 204(1):19 e1–6, 2011.
- [235] J. S. Kwon, A. Tinker, G. Pansegrau, J. McAlpine, M. Housty, M. McCullum, and C. B. Gilks. Prophylactic salpingectomy and delayed oophorectomy as an alternative for BRCA mutation carriers. *Obstet Gynecol*, 121(1):14–24, 2013.
- [236] S. A. Narod. Salpingectomy to prevent ovarian cancer: A Countercurrents series. *Curr Oncol*, 20(3):145–7, 2013.
- [237] R. Manchanda and U. Menon. Opportunistic bilateral salpingectomy (OBS) for the prevention of ovarian cancer should be offered in the context of a clinical trial: For: There is lack of clarity on a number of key issues. *BJOG*, 123(3):463, 2016.
- [238] M. G. Harmsen, M. Arts-de Jong, N. Hoogerbrugge, A. H. Maas, J. B. Prins, J. Bulten, S. Teerenstra, E. M. Adang, J. M. Piek, H. C. van Doorn, M. van Beurden, M. J. Mourits, R. P. Zweemer, K. N. Gaarenstroom, B. F. Slangen, M. C. Vos, L. R. van Lonkhuijzen, L. F. Massuger, R. P. Hermens, and J. A. de Hullu. Early salpingectomy (TUbectomy) with delayed oophorectomy to improve quality of life as alternative for risk-reducing salpingo-oophorectomy in BRCA1/2 mutation carriers (TUBA study): a prospective non-randomised multicentre study. *BMC Cancer*, 15: 593, 2015.
- [239] R. Hadfield, H. Mardon, D. Barlow, and S. Kennedy. Delay in the diagnosis of endometriosis: a survey of women from the USA and the UK. *Hum Reprod*, 11(4):878–80, 1996.
- [240] W. Kuohung, G. L. Jones, A. F. Vitonis, D. W. Cramer, S. H. Kennedy, D. Thomas, and M. D. Hornstein. Characteristics of patients with endometriosis in the United States and the United Kingdom. *Fertil Steril*, 78(4):767–72, 2002.
- [241] B. J. Cairns, B. Liu, S. Clennell, R. Cooper, G. K. Reeves, V. Beral, and D. Kuh. Lifetime body size and reproductive factors: comparisons of data recorded prospectively with self reports in middle age. *BMC Med Res Methodol*, 11:7, 2011.
- [242] F. L. Wright, J. Green, G. Reeves, V. Beral, B. J. Cairns, and Million Women Study Collaborators. Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. *BMC Med Res Methodol*, 15:81, 2015.
- [243] Associated Press. Johnson and Johnson to pay \$72m in case linking baby powder to ovarian cancer. *The Guardian*, 24 Feb, 2016.
- [244] Reuters in New York. Johnson and Johnson loses another court case over talcum powder and cancer. *The Guardian*, 3 May, 2016.

APPENDIX A

MILLION WOMEN STUDY RECRUITMENT QUESTIONNAIRE

This questionnaire was completed by study participants at recruitment, in 1996-2001. I had no role in developing the questionnaire.

THE MILLION WOMEN STUDY

A national survey of women invited for breast screening

We need one million women to help us in research that will benefit women all over the world.

Would you become one of these special women?

More and more women are taking hormone replacement therapy (HRT) so it is vital that we find out as much as possible about its benefits and any possible side effects. We have a unique opportunity through the NHS Breast Screening Programme to learn about the way different types of HRT and other lifestyle factors affect a woman's health, particularly her breasts. Britain is the only country in the world that can carry out this study because it is the only one with the combination of a large population and a comprehensive national breast screening programme.

The NHS Breast Screening Programme, the Imperial Cancer Research Fund and the Medical Research Council have joined together to organise The Million Women Study. If one million women answer this questionnaire over the next three years we could have some of the answers to our most important questions about HRT within five years or so.

We would be very grateful if you could set aside some time to answer these questions. It should not take more than 10-15 minutes. You do not have to answer this questionnaire and if you decide not to you will still have your screening done in the normal way.

Please answer every question and do not leave blanks as all the information that you give us is very useful. If you are not sure about exact dates or ages an approximate answer is better than none. If you have any questions you can ring us on freephone **0800 262 872**.

Even if you are not taking HRT it is just as important that you fill in the questionnaire.

Please bring this questionnaire to your breast screening appointment

To help us read your answers please write as clearly as possible and be sure to complete the questionnaire as shown:

Please put a cross in the appropriate box(es) ☒ ☐ ☐

OR put numbers in the appropriate box e.g. 23rd April 1946 , , age years

GENERAL QUESTIONS ABOUT YOU

1. What is your date of birth? (please put day/month/year)

//

2. How old are you? years

3. How tall are you? (please give to the nearest inch)

 feet inches

4. About how much do you weigh?

 stone lbs

5. How old were you when you finished full time schooling? (please cross one box)

☐ did not go to school	☐ 15
☐ 13 or younger	☐ 16
☐ 14	☐ 17 or older

6. What qualification(s) do you have from school, college or the equivalent?

(please put a cross in the most appropriate box(es))

☐ clerical or commercial qualifications (eg secretarial, hairdressing etc)
☐ nursing or teaching
☐ "O" level (or equivalent)
☐ "A" level (or equivalent)
☐ college/university degree (or equivalent)
☐ none of these

7. About how many cigarettes do you smoke on average each day, now? (please cross one box)

☐ none	☐ 15-19
☐ less than 5	☐ 20-24
☐ 5-9	☐ 25 or more
☐ 10-14	

8. Are you an ex-smoker? ☐ No ☐ Yes

9. About how much wine, beer or spirits do you drink on average each week? (please cross one box for each type)

Wine (glasses per week)	Lager/Cider/Beer (half pints per week)	Spirits (tots per week)
☐ none	☐ none	☐ none
☐ less than 1	☐ less than 1	☐ less than 1
☐ 1-3	☐ 1-3	☐ 1-3
☐ 4-6	☐ 4-6	☐ 4-6
☐ 7-10	☐ 7-10	☐ 7-10
☐ 11-15	☐ 11-15	☐ 11-15
☐ 16-20	☐ 16-20	☐ 16-20
☐ 21+	☐ 21+	☐ 21+

If you drink wine is it

☐ mostly red	☐ mostly white
☐ about the same amount of red and white?	

10. How often do you do any exercise?

- ☐ rarely/never ☐ 2-3 times a week
☐ less than once a week ☐ 4-6 times a week
☐ once a week ☐ every day

11. How often do you do strenuous exercise?

(that is, enough to cause sweating or a fast heart beat.)

- ☐ rarely/never ☐ 2-3 times a week
☐ less than once a week ☐ 4-6 times a week
☐ once a week ☐ every day

QUESTIONS ABOUT YOU AND YOUR FAMILY

12. Have you ever had any children? ☐ No ☐ Yes

- if No, please go on to question 15

13. How many children have you had?

(please include stillbirths; it is not necessary to include miscarriages)

14. When was each child born, and for how many months did you breastfeed each child, if at all?

DATE OF BIRTH

(If you had twins or triplets please repeat the same date for each child)

BREASTFEEDING

(Months that you breastfed each child; put "0" if you did not breastfeed that child and "1" if you breastfed for a month or less)

	day	month	year	months
1st child				
2nd child				
3rd child				
4th child				
5th child				
6th child				
7th child				
8th child				
9th child				
10th child				

15. Have you ever been for breast screening before?

- ☐ No
☐ Yes- If Yes, about how many years ago was your last screen? years ago

16. Have you ever had a breast lump removed or any operations on your breast(s)?

- ☐ No
☐ Yes- If Yes, how old were you? years

(If you have had more than one operation please write your age at the first operation)

17. Have you ever had breast cancer diagnosed?

- ☐ No
☐ Yes- If Yes, how old were you when the cancer was first diagnosed? years

18. Has your mother ever had breast cancer diagnosed?

- ☐ No ☐ Don't know
☐ Yes- If Yes, how old was she when the cancer was first diagnosed? years

19. How many sisters do you have? sisters

(put "0" if you do not have any sisters, please include any sisters who have died)

20. Have any of your sisters ever had breast cancer diagnosed?

- ☐ No/No sisters ☐ Don't know
☐ Yes- If Yes, how old were they when the cancer was first diagnosed?

1st sister years 2nd sister years

QUESTIONS ABOUT YOUR HEALTH

21. Have you ever had any other cancer?

- ☐ Yes ☐ No

Please describe

22. Have you EVER had:

(please cross "Yes" or "No" for each condition)

- High blood pressure - when pregnant ☐ Yes ☐ No
 High blood pressure - when not pregnant ☐ Yes ☐ No
 Heart disease(eg heart attack/angina) ☐ Yes ☐ No
 Stroke ☐ Yes ☐ No
 Diabetes ☐ Yes ☐ No
 High blood cholesterol ☐ Yes ☐ No
 Blood clot (thrombosis) ☐ Yes ☐ No

23. Are you NOW being treated for:

- High blood pressure (hypertension) ☐ Yes ☐ No
 Heart disease ☐ Yes ☐ No
 Diabetes ☐ Yes ☐ No
 High blood cholesterol ☐ Yes ☐ No
 Varicose veins ☐ Yes ☐ No
 Clotting problems ☐ Yes ☐ No
 Asthma ☐ Yes ☐ No
 Rheumatoid arthritis ☐ Yes ☐ No
 Osteoarthritis ☐ Yes ☐ No
 Thyroid problems ☐ Yes ☐ No
 Osteoporosis ☐ Yes ☐ No
 Depression/Anxiety ☐ Yes ☐ No

24. Are you **NOW** being treated for any other **serious** illness?

☐ Yes ☐ No

Please describe this illness

Please describe the treatment

QUESTIONS ABOUT PAST OPERATIONS

25. Have you had a hysterectomy?

☐ No
☐ Yes- If **Yes**, how old were you? years

26. Have you had **BOTH** ovaries removed?

☐ No ☐ Not sure
☐ Yes- If **Yes**, how old were you? years

27. Have you been sterilised (had your tubes tied)?

☐ No
☐ Yes- If **Yes**, how old were you? years

QUESTIONS ABOUT YOUR USE OF THE PILL

28. Have you ever used the pill (oral contraceptive)?

☐ Yes
☐ No - if **No**, please go to question 32

29. About how old were you when you **first went on** the pill? years

30. About how old were you when you **last came off** the pill? years

31. For how many years **in total** did you take the pill?

years

(Add together the years and months when you actually took the pill -do not count the years and months when you were not taking it. Please write "0" if you used the pill for less than a year in total)

QUESTIONS ABOUT YOUR USE OF HORMONE REPLACEMENT THERAPY (HRT)

32. Have you ever used hormone replacement therapy (HRT)? ☐ No - if **No** - please go to question 39

☐ Yes

33. How old were you when you **first started** using HRT?

years

34. Had your periods stopped before you started using HRT? (Cross "Yes" if you had a hysterectomy before starting HRT)

☐ No
☐ Yes - if **Yes**, how old were you when your periods stopped? years

35. For about how many years **in total** have you used HRT? years

(Add together the years and months when you used HRT - do not count the years and months when you were not using HRT. Please write "0" if you used HRT for less than a year in total)

36. Are you **now** using HRT?

☐ Yes
☐ No - if **No**, how old were you when you last used HRT? years

37. What is the name of the most **RECENT** HRT you have used?

☐ Prempak C 0.625mg	☐ Premarin 0.625mg
☐ Prempak C 1.25mg	☐ Premarin 1.25mg
☐ Tridestra	☐ Evorel 25mcg/50mcg
☐ Trisequens	☐ Evorel 75mcg/100mcg
☐ Trisequens Forte	☐ Progynova 1mg
☐ Cycloprogynova 1mg	☐ Progynova 2mg
☐ Cycloprogynova 2mg	☐ Estraderm 25mcg
☐ Estrapak	☐ Estraderm 50mcg
☐ Estracombi	☐ Estraderm 100mcg
☐ Climaval 1mg	☐ Zumenon 1mg
☐ Climaval 2mg	☐ Zumenon 2mg
☐ Premique Cycle	☐ Ethinyloestradiol
☐ Premique	☐ Micronor
☐ Nuvelle	☐ Provera
☐ Kliofem	☐ Duphaston
☐ Livial	
☐ Do not know	☐ Implants
	☐ Oestrogel

Other (please write here)

38. For how many years **in total** did you use the most recent type of HRT? years

(Please write "0" if you used this recent HRT for less than a year in total)

QUESTIONS ABOUT YOUR PERIODS

39. About how old were you when your periods started? years

40. Have your periods NOW stopped?

Cross "Yes"-if you are not having periods now, either because of your menopause, after a hysterectomy or after stopping HRT.

Cross "No"- if you are still having regular periods now, even if they are because you are taking HRT.

Cross "Irregular"-if your periods have been irregular and you think it might be because of the menopause.

☐ Yes- If Yes, how old were you when they stopped? years

☐ No

☐ Irregular

FINAL SECTION

41. So that we can find out about your health in the future we may need to contact you again or look at your screening or medical records. We would be grateful if you gave us permission to contact you again or to use information from those records.

We guarantee that all information obtained will be treated with absolute confidentiality and used for medical research only. Of course, you do not have to give permission. Your response to this request will not affect your screening or the treatment you receive in any way.

If you give permission, please sign here and print your name, address and other details in the section below. Please print in BLOCK CAPITALS as clearly as possible.

Signature:

today's date:

Surname:

Given name(s):

House number and street:

District:

Town:

County:

Postcode:

Surname at birth:

Town of birth:

NHS (National Health Service) Number:
(The number on your medical card)

Breast Screening Number:
(This is in the top left hand corner of your **screening invitation letter** and starts with the letters LGL)

Please bring this questionnaire to your breast screening appointment

If you would like to post this questionnaire back to us, please send it to:

**THE MILLION WOMEN STUDY CO-ORDINATING CENTRE
ICRF-CEU, GIBSON BUILDING,
RADCLIFFE INFIRMARY,
OXFORD OX2 6HE**

FREEPHONE: 0800 262 872

**Breast Screening Unit
Linton House
Thirlestaine Road
Cheltenham
GL53 7AS**

LGL

☐

THANK YOU VERY MUCH FOR YOUR HELP

APPENDIX B

ETHICAL APPROVAL FOR THE MWS OVARIAN CANCER TISSUE PILOT STUDY

B.1 MWS APPENDIX 28: OVARIAN TUMOUR TISSUE SUBSTUDY

I drafted this appendix to the Million Women Study Protocol, in collaboration with others (particularly Professor Gillian Reeves and Professor Jane Green).

THE MILLION WOMEN STUDY

A national study of women's health

MREC application reference no. 97/5/01

APPENDIX TO PROTOCOL: OVARIAN TUMOUR TISSUE SUBSTUDY

Part 1. Feasibility Study

Aim

We plan to utilise stored archival tumour tissue from women diagnosed with ovarian cancer within the Million Women Study cohort with the aim of examining the aetiology of molecular subtypes of ovarian cancer. In order to assess the feasibility of large-scale tumour tissue analyses we plan to conduct a pilot study among ovarian cancer cases treated within the Oxford University Hospitals Trust. The results of this study will be used to inform the development of a protocol for large-scale analyses of ovarian tumour tissue within the study.

Background and rationale

Although the Million Women Study was initially set up primarily to describe the role of hormonal and other risk factors in breast cancer incidence, the study also provides unparalleled opportunities to explore relationships between such risk factors and other diseases. This is particularly the case for ovarian cancer, as it shares many biological similarities to breast cancer (in being at least partly hormonally-driven), but is difficult to study prospectively in any but the largest studies, due to its relative rarity.

Over the last decade or so, findings from the Million Women Study¹⁻⁵ and other epidemiological studies have shown that ovarian cancer is a heterogeneous disease, and that the impact of many established risk factors varies between the histological tumour types (histotypes). Additionally, molecular and histopathological studies suggest that the different histotypes may have very distinct, and possibly extra-ovarian, origins (for example, it is hypothesised that some serous tumours may originate in the fallopian tubes, while some endometrioid and clear cell tumours may arise from endometriosis).⁶ Although the clinical diagnosis of ovarian cancer histotype is still predominantly on the basis of morphology, the number of histotypes has changed in recent years (for example, high-grade and low-grade serous tumours are now thought to be distinct tumour types, with different origins), leading to a different classification system to that used in historically-diagnosed samples.⁷ It has also been suggested that subtyping of ovarian cancer may be done more reliably with the aid of immunohistochemical markers.⁸

Although routine data on a limited number of tumour characteristics has recently become available through record linkage to NHS datasets, analysis of tumour tissue is required to obtain detailed information about tumour pathology in order to classify tumours into relevant molecular subtypes.

Access to such data would substantially enhance the scientific value of the cohort as it would enable us to examine how specific molecular subtypes of ovarian cancer vary in terms of risk factors and survival. This work will not only provide reliable information on risk factors for rarer forms of ovarian cancer but should also help clarify the biological mechanisms underlying the development of the disease.

Since processing and analysis of tumour tissue is both expensive and labour intensive, it is important to identify the most reliable and efficient means of obtaining relevant information on tumour pathology from relatively large numbers of tissue samples. For this reason, we first plan to conduct a feasibility study among those Million Women Study participants treated for ovarian cancer within the Oxford University Hospitals Trust. The results of this study will be used to inform a larger study of tissue collection within the cohort.

We have previously sought (and been granted) ethics approval to conduct a similar feasibility study of retrieving archival samples of tumour tissue from Million Women Study participants with breast cancer (MWS Protocol Appendix V1.1 2nd March 2015, approved 01/04/2015). The present proposal is based on that for breast cancer tumour tissue, and the overarching principles are the same. However, we believe that a separate feasibility study is necessary for ovarian cancer tumour tissue, as the nature of the original samples collected may be different, and the specific tests that would be required for molecular characterisation will be different.

Study design

The feasibility study aims to identify, process and analyse tumour tissue from the records of all Million Women Study participants diagnosed with incident ovarian cancer (that is after recruitment into the study) and subsequently treated for ovarian cancer at one or more hospitals within the NHS Oxford University Hospitals Trust. Our records suggest that around 150 women with invasive ovarian cancer will be eligible for inclusion into the study. The process by which relevant ovarian cancer cases will be identified and samples processed is summarised in Figure 1.

Study population

We will identify women with ovarian cancer who we believe were treated in Oxford hospital(s), by using the hospital code as recorded in data we already hold from Public Health England (obtained via the Health and Social Care Information Centre, HSCIC). Women eligible for the feasibility study will be those who have had ovarian cancer diagnosed following recruitment into the Million Women Study and who are identified as having ovarian tumour tissue stored at the NHS Oxford University Hospitals Trust pathology laboratory.

Biological samples: retrieval, processing and storage

Retrieval and on-site processing of the tumour tissue samples will be carried out at the Oxford Centre for Histopathology Research (OCHRe), an integrated service-driven facility which enables access to equipment and methodologies for research. The OCHRe is also closely linked to the Oxford Radcliffe

Biobank (ORB) which provides a simple and efficient way to collect and store samples according to regulatory requirements.

We will send the OCHRe an encrypted datafile containing a list of women who we believe should be on their database, as having been treated for ovarian cancer in Oxford. The datafile will contain three variables: Date of birth (DOB), NHS number, and a unique identifier generated just for this study. The Million Women Study ID number, under which participant information is held in our secure study database, will not be shared and the key linking the unique study ID with the MWS ID will be held separately from the MWS database. Administration staff at the OCHRe will use the NHS number and DOB to link to their database, add information from the pathology record, and return the enhanced datafile. For sample processing by laboratory staff (see below), cases will be identified by means of the unique study identifier only.

For each eligible woman, the OCHRe will provide the study investigators with a copy of the corresponding pathology report, which documents the types of tissue sample(s) available and which histopathological tests have already been carried out, and, where possible, retrieve formalin-fixed paraffin blocks of tissue (which may include samples of the ovary, metastatic tumour, and fallopian tubes). From these blocks, sections and tissue cores will be taken for the purposes of immunohistochemistry and molecular studies and, where appropriate, tissue micro-arrays will be prepared. Samples will be tested for a range of immunohistochemical markers, which may include (but is not limited to): ARID1A, CDKN2A, DKK1, HNF1B, MDM2, PGR, TP53, TFF3, VIM, WT1, cyclin E1, and NCAPH. The number of blocks retrieved, and the type of sections or cores taken, will depend on the nature and amount of tissue available with a view to avoiding unnecessary depletion of this valuable resource. Where sufficient tumour tissue exists, additional samples will be taken and stored for future studies for which separate ethics approval will be sought.

Immunohistochemical staining of slides will make use of a work pipeline developed by Professor Ahmed, Director of the Ovarian Cancer Cell Laboratory at the Weatherall Institute of Molecular Medicine, and Professor of Gynaecological Oncology at the University of Oxford. Professor Ahmed has previously worked in collaboration with OCHRe on projects involving tissue samples from ovarian cancer patients in Oxford. His research group has developed expertise in quality control assessment of novel antibodies, and works in collaboration with GCP Laboratories (based in the Department of Oncology at the University of Oxford) to ensure standardised and validated immunohistochemical staining. Prof Ahmed's laboratory and GCP Laboratories have approval (ethics committee number: 11/SC/0014) to process tissue and store anonymised samples from ovarian cancer patients, including those originating from collaborating groups.

All eligible tissue samples, including slides, sections and cores that are taken for the purposes of the pilot study, will remain within the Oxford University Hospitals Trust, or within the approved laboratories mentioned above. Information on tumour markers derived from the samples will be provided to the study investigators by OCHRe via the unique identification number

generated solely for the purposes of this study. These data will then be entered onto a secure database at the Million Women Study co-ordinating centre for further analysis.

Analysis

Since this is intended to be a feasibility study it is not expected that this relatively small subset of data will be sufficient to reliably address any of the scientific questions included in the rationale for the study. However, these pilot data will be analysed to provide preliminary descriptive findings on the distribution of key histopathological tumour markers according to age, other tumour characteristics and genetic and environmental factors. These findings will help to inform estimates of statistical power in any subsequent large-scale study. They may also be used for comparison with/ validation of histopathological diagnoses received from national cancer registries.

Consent, confidentiality and ethical issues

All women taking part in the Million Women Study (and hence this substudy) have given signed consent for follow-up and access to their medical record. The substudy does not require any further contact with participants as information on ovarian cancer diagnosis is already routinely provided to the study investigators and the availability of an appropriate tissue sample can be ascertained via linkage to their medical record.

The substudy will conform to current ethical and legal guidelines regarding consent, confidentiality and the use of human biological samples. It will be conducted in accordance with relevant aspects of the Data Protection Act and the Council of Europe's Recommendation on the Protection of Medical Data. It will follow the guidance outlined in the Medical Research Council's documents on Personal Information in Medical Research and on Collections of Human Tissue and Biological Samples for Use in Medical Research. Each biological sample will be physically identified only by a study-specific code number. The information needed to link this number with questionnaire data, which may include personal identifying data where necessary for future follow-up within the Million Women Study, will be held securely and separately from other databases. All data will be analysed only in anonymised form and future publications will not identify individual women taking part. Our detailed data security policies are available online at www.ceu.ox.ac.uk/policies2.

Publication of results

The Million Women Study is a joint research project of the NHS Breast Screening Programme, the Medical Research Council and Cancer Research UK. Any results from this substudy will be published in peer-reviewed journals and where appropriate the relevant authorities (e.g. the Department of Health) will be notified of the findings.

References

1. Beral V, Doll R, Hermon C, Peto R, Reeves G. Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies

including 23,257 women with ovarian cancer and 87,303 controls. *Lancet* 2008;**371**: 303-14.

2. Beral V, Gaitskell K, Hermon C, Moser K, Reeves G, Peto R. Ovarian cancer and smoking: individual participant meta-analysis including 28,114 women with ovarian cancer from 51 epidemiological studies. *Lancet Oncol* 2012;**13**: 946-56.

3. Collaborative Group On Epidemiological Studies Of Ovarian Cancer. Ovarian cancer and body size: individual participant meta-analysis including 25,157 women with ovarian cancer from 47 epidemiological studies. *PLoS Med* 2012;**9**: e1001200.

4. Collaborative Group On Epidemiological Studies Of Ovarian Cancer. Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. *Lancet* 2015;**385**(9980):1835-1842.

5. Gaitskell K, Green J, Pirie K, Reeves G, Beral V, on behalf of the Million Women Study Collaborators. Tubal ligation and ovarian cancer risk in a large cohort: Substantial variation by histological type. *Int J Cancer* 2016;**138**(5):1076-1084.

6. Kurman RJ, Shih IeM. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer--shifting the paradigm. *Hum Pathol* 2011;**42**: 918-31.

7. Kurman RJ, Carcangiu ML, Herrington CS, Young RH, (Eds.). *WHO Classification of Tumours of Female Reproductive Organs*, 4th ed. Lyon: IARC, 2014.

8. Kobel M, Kalloger SE, Lee S, Duggan MA, Kelemen LE, Prentice L, Kalli KR, Fridley BL, Visscher DW, Keeney GL, Vierkant RA, Cunningham JM, et al. Biomarker-based ovarian carcinoma typing: a histologic investigation in the ovarian tumor tissue analysis consortium. *Cancer Epidemiol Biomarkers Prev* 2013;**22**: 1677-86.

Part 2. Linkage to existing studies of ovarian cancer in Oxford

Aim

We would like to link the Million Women Study cohort with an existing database of clinical studies of women with ovarian cancer who have been treated in Oxford, co-ordinated by Prof Ahmed Ahmed, Professor of Gynaecological Oncology at the University of Oxford. The aim is to identify if any participants in the Million Women Study are also participants in Prof Ahmed's studies. If so, then collaboration between the research groups would potentially enhance the value of both studies, and allow for more efficient use of resources.

Background and rationale

Professor Ahmed is conducting several clinical studies of patients with ovarian cancer in Oxford. As described above, we wish to collaborate with Prof Ahmed as part of the ovarian tumour tissue substudy, so that we can utilise his expertise and existing work pipelines.

In addition, it is possible that at least some of the participants in the Million Women Study who have been treated for ovarian cancer in the Oxford University Hospitals Trust may have also been recruited to participate in other studies of ovarian cancer being run by Prof Ahmed. If so, then these women may have also had additional samples taken for these other studies (such as blood, frozen tumour tissue, or other material), some of which may already have been processed and analysed. We thus wish to identify if any of the women participating in Prof Ahmed's studies are also participants in the Million Women Study, so that we can collaborate and make the best use of all the information and samples available.

Study design

This proposed linkage is initially primarily exploratory, to see how many women (if any) are participants in both the Million Women Study and Prof Ahmed's clinical studies, and what additional clinical information and/or biological specimens are available. Further specific studies may then be planned, depending on what resources are identified.

Study population

The linkage study will consist of women who are participants in the Million Women Study, and are also participants in Prof Ahmed's clinical studies of ovarian cancer. This will be a (small) subset of women who are participants in the Million Women Study and have been treated for ovarian cancer in Oxford.

As with the ovarian tumour tissue substudy (see Part 1, above), we will identify women with ovarian cancer who we believe were treated in Oxford hospital(s), by using the hospital code as recorded in data we already hold from Public Health England (obtained via the Health and Social Care Information Centre, HSCIC).

Linkage

This will use a similar process to that used for linkage to the OCHRe database (described above). We will send a designated member of Prof Ahmed's research team an encrypted datafile containing a list of women who we believe may be on their database, as having been treated for ovarian cancer in Oxford. The datafile will contain three variables: Date of birth (DOB), NHS number, and a unique identifier generated just for this linkage study. This will allow the designated member of Prof Ahmed's research team to use the NHS number and DOB to link to their database.

Initially, the aim of the linkage will be to identify a) the number of women (if any) who are participants in both studies, and b) the additional clinical information and biological samples available. This information will be added, and the datafile returned to us; we will then merge this datafile with our whole dataset, using the unique identifier. We will not share the Million Women Study ID number.

Consent, confidentiality and ethical issues

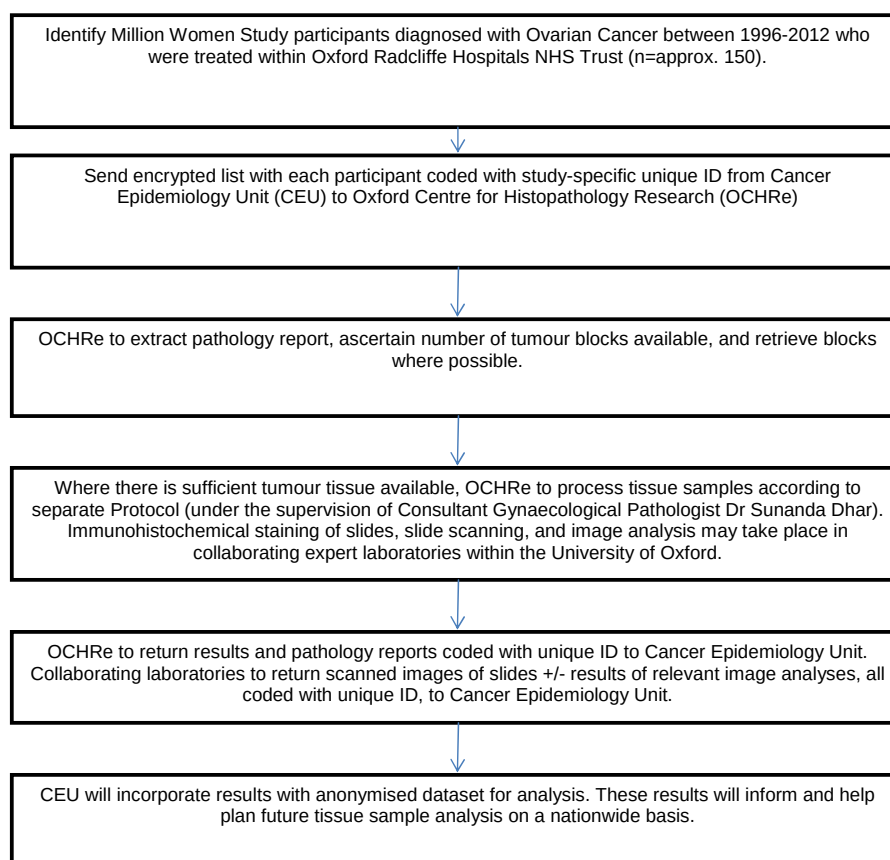
All women taking part in the Million Women Study and in Prof Ahmed's clinical studies of ovarian cancer (and hence this proposed linkage substudy) have given signed consent for follow-up and access to their medical record. We have previously obtained ethical approval for linkage between the Million Women Study and two other UK medical research studies, the MRC National Survey of Health and Development Cohort /1946 Birth Cohort (approved 15/01/2008) and UK Biobank (approved 10/07/2012).

The proposed linkage study will conform to current ethical and legal guidelines regarding consent, confidentiality and the use of human biological samples. It will be conducted in accordance with relevant aspects of the Data Protection Act and the Council of Europe's Recommendation on the Protection of Medical Data. It will follow the guidance outlined in the Medical Research Council's documents on Personal Information in Medical Research and on Collections of Human Tissue and Biological Samples for Use in Medical Research. The substudy does not require any further contact with participants as information on ovarian cancer diagnosis is already routinely provided to the study investigators and their participation in Prof Ahmed's studies can be ascertained via linkage to his secure study database.

Publication of results

The Million Women Study is a joint research project of the NHS Breast Screening Programme, the Medical Research Council and Cancer Research UK. Professor Ahmed's studies are jointly funded by Ovarian Cancer Action, Target Ovarian Cancer, and the Medical Research Council. Any results from this substudy will be published in peer-reviewed journals and where appropriate the relevant authorities (e.g. the Department of Health) will be notified of the findings. However, as the initial aim of this proposed linkage substudy is primarily exploratory, publishable results may not be produced.

Figure 1. Flow Chart describing Pilot Study to assess feasibility of Ovarian Tissue Sample processing and analyses for participants in The Million Women Study



APPENDIX B. ETHICAL APPROVAL FOR THE MWS OVARIAN
CANCER TISSUE PILOT STUDY

**B.2 LETTER CONFIRMING ETHICAL APPROVAL FOR THE MWS OVARIAN
TUMOUR TISSUE PILOT STUDY**

East of England - Cambridge South Research Ethics Committee

The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

04 March 2016

Professor Valerie Beral
Director, Cancer Research UK Epidemiology Unit, Oxford
University of Oxford
Richard Doll Building
Roosevelt Drive
Oxford
OX3 7LF

Dear Professor Beral

Study title:	The Million Women Study
REC reference:	97/5/001
Amendment number:	Amendment 28 (substantial)
Amendment date:	24 February 2016
Study title:	The Million Women Study

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Covering letter on headed paper [From Lynden Guiver, University of Oxford]	Amendment 28 (substantial)	16 February 2016
Notice of Substantial Amendment (non-CTIMP) [Signed by Professor Beral and Dr Ronja Bahadori, University of Oxford]	Amendment 28 (substantial)	24 February 2016
Other [MWS Appendix to Protocol]	1.2	24 February 2016

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at <http://www.hra.nhs.uk/hra-training/>

97/5/001:	Please quote this number on all correspondence
------------------	---

Yours sincerely

A handwritten signature in black ink, appearing to read 'PP: [Signature]', is enclosed in a rectangular box.

Dr Leslie Gelling
Chair

E-mail: nrescommittee.eastofengland-cambridgesouth@nhs.net

Enclosures: List of names and professions of members who took part in the review

East of England - Cambridge South Research Ethics Committee

Attendance at Sub-Committee of the REC meeting on 04 March 2016

Committee Members:

<i>Name</i>	<i>Profession</i>	<i>Present</i>	<i>Notes</i>
Dr Leslie Gelling	(Chair) Reader in Research Ethics	Yes	
Dr Frank Wells	(Vice-Chair) Retired Pharmaceutical Physician	Yes	

Also in attendance:

<i>Name</i>	<i>Position (or reason for attending)</i>
Nicola Kohut	REC Assistant

APPENDIX C

PAPERS PUBLISHED FROM DPHIL WORK

C.1 TUBAL LIGATION AND OVARIAN CANCER RISK IN A LARGE COHORT: SUBSTANTIAL VARIATION BY HISTOLOGICAL TYPE (INTERNATIONAL JOURNAL OF CANCER, 2016)

I am the first author and local guarantor for this paper, which is based on work in Chapter 5. I conducted the statistical analyses for this project, drafted the paper, and revised it in response to reviewers' comments, with input from co-authors.

Tubal ligation and ovarian cancer risk in a large cohort: Substantial variation by histological type

Kezia Gaitskell, Jane Green, Kirstin Pirie, Gillian Reeves and Valerie Beral;
on behalf of the Million Women Study Collaborators

Cancer Epidemiology Unit, University of Oxford, Oxford, United Kingdom

Histopathological and molecular studies suggest that different histological subtypes (histotypes) of ovarian cancer have different aetiologies. Few studies have been large enough to explore reliably the effect of tubal ligation (sterilization), which has been associated with a reduced overall risk of ovarian cancer, on different tumour histotypes. In a prospective study of 1.1 million UK women without prior cancer or bilateral oophorectomy, 8,035 ovarian cancers occurred during mean follow-up of 13.8 years. Using a Cox proportional hazards model, we estimated adjusted relative risks of ovarian cancer associated with tubal ligation. Overall, there was substantial heterogeneity in tumour risk associated with tubal ligation for the four main histotypes, serous, endometrioid, mucinous and clear cell (heterogeneity: $p < 0.0001$). For serous tumours, the most common histotype ($n = 3,515$), risks differed significantly between high-grade (RR: 0.77, 95% CI: 0.67–0.89) and low-grade tumours (RR: 1.13, 95% CI: 0.89–1.42); heterogeneity: $p = 0.007$. Relative risks were almost halved for endometrioid ($n = 690$, RR: 0.54, 95% CI: 0.43–0.69) and clear cell tumours ($n = 401$, RR: 0.55, 95% CI: 0.39–0.77), but there was no association between tubal ligation and mucinous tumours ($n = 836$, RR: 0.99, 95% CI: 0.84–1.18). For the main tumour histotypes we found little variation of risk by timing of tubal ligation. The significant differences by tumour histotype are unlikely to be due to confounding and are consistent with hypotheses that high-grade and low-grade serous tumours have different origins, and that some endometrioid and clear cell tumours might arise from cells and/or carcinogens travelling through the fallopian tubes.

Ovarian cancer is one of the most common causes of cancer death amongst women in high income countries.¹ There is increasing evidence that the different histological subtypes (histotypes) of ovarian cancer may have different origins, and

show different molecular signatures.^{2–5} The vast majority of ovarian cancers are epithelial tumours, of which the four most common histotypes are serous, mucinous, endometrioid and clear cell tumours.⁶ For serous tumours, it has been further suggested that they be divided into low-grade serous tumours (consisting of serous borderline tumours and low-grade serous carcinoma) and high-grade serous carcinoma.

Meta-analyses of observational epidemiological studies have consistently found that tubal ligation (also known as sterilization, in which the fallopian tubes are clipped, cut, or tied) is associated with an overall reduced risk of ovarian cancer. However, published findings by tumour histotype are inconsistent, except perhaps that most investigators have reported a reduced risk for endometrioid tumours.^{7–9} Epidemiological studies need to be sufficiently large to have enough power to detect reliably heterogeneity by tumour histotype. We therefore report here on the association observed between tubal ligation and incident ovarian cancer in a large cohort study of UK women.

Key words: ovarian cancer, tubal ligation, sterilization, histological subtypes, histotype

Abbreviations: BMI: body mass index; CI: confidence interval; FIGO: International Federation of Gynecology and Obstetrics; ICD: International Classification of Diseases; IQR: interquartile range; NCIN: National Cancer Intelligence Network; NHS: National Health Service; NHSCR: National Health Service central register; RR: relative risk; SD: standard deviation; WHO: World Health Organization

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Grant sponsor: Cancer Research UK; **Grant number:** C570/A16491, C38302/A17318; **Grant sponsor:** Medical Research Council; **Grant number:** MR/K02700X/1

DOI: 10.1002/ijc.29856

History: Received 13 July 2015; Accepted 1 Sep 2015; Online 17 Sep 2015

Correspondence to: Dr Kezia Gaitskell, Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Richard Doll Building, Roosevelt Drive, Oxford OX3 7LF, UK, Tel.: +44-1865-289651, Fax: +44-1865-289610, E-mail: kezia.gaitskell@ceu.ox.ac.uk

Material and Methods

Study design, data collection and follow-up

The Million Women Study is a prospective study of 1.3 million UK women, recruited in 1996–2001 via the UK National Health Service (NHS) Breast Screening Programme. At recruitment, women completed a questionnaire on socio-demographic, reproductive, medical and lifestyle factors. The cohort is resurveyed approximately every 3 to 5 years. The

What's new?

Tubal ligations can reduce ovarian cancer—but possibly only certain types. This large study reports on the risk of four main ovarian tumour histotypes following tubal ligation. They found that the procedure almost halved the risk of clear cell tumors and endometrioid tumours, but had little effect on mucinous tumours. For serous tumours, risk varied by grade: tubal ligation reduced the risk of high-grade, but not low-grade, serous tumours. These findings support the hypothesis that high- and low-grade serous tumours have different causes, and that some ovarian tumours may arise from cells or carcinogens travelling through the fallopian tubes.

study design and methods are described in detail elsewhere,^{10,11} and questionnaires can be viewed online at <http://www.millionwomenstudy.org>.

All participants have been flagged on the NHS Central Register (NHSCR), so the study investigators are routinely notified of cancer registrations and deaths. The information provided includes the date of the event (cancer registration or death), together with the cancer site (coded using the 10th revision of the International Classification of Diseases, ICD-10)¹² and tumour morphology (coded using the second and third editions of the International Classification of Diseases for Oncology, ICD-O).^{13,14} The Million Women Study has also been linked to data from the National Cancer Intelligence Network (NCIN), providing additional information on diagnostic histology codes, and on tumour grade, for cases diagnosed in England. For this study, tumour histology and grade information from NCIN were used where available, to supplement the NHSCR data.

All participants gave written consent to follow-up at recruitment. Ethical approval was granted by the Oxford and Anglia Multi-Centre Research Ethics Committee (MREC 97/01).

Exposure variables

Women in the study were asked at recruitment “Have you been sterilised (had your tubes tied)?,” and if so, their age at sterilization (used for analyses of the timing of sterilization).

Outcome

The outcome of interest was ovarian cancer (C56 in ICD-10).¹² For histotype analyses, the outcome was split into five histological groups: serous (ICD-O codes 8441–8442, 8451, 8460–8463, 9014), mucinous (ICD-O codes 8470–8490), endometrioid (ICD-O codes 8380–8381, 8560, 8570, 8933, 8950), clear cell (ICD-O codes 8310, 8313) and other.

For some exploratory analyses, tumour types were analysed separately by grade (available for about 40% of all cases). Three-tier grading systems had generally been used (including those outlined by FIGO (International Federation of Gynecology and Obstetrics),¹⁵ WHO (World Health Organization),¹⁶ and Silverberg¹⁷), and the few cases classified as “grade 4” were grouped together with “grade 3” tumours.

Serous tumours were divided into low-grade tumours [defined here as borderline (ICD-O codes 8442, 8451, 8462, 8463) or grade 1 serous tumours], and high-grade serous carcinoma (defined here as grade ≥ 2 serous tumours). The

assumptions we made in converting tumour grade information from the three-tier (grade 1–3) to the two-tier (low-grade *vs.* high-grade) system for serous carcinoma were in line with those used in other publications.^{4,18,19}

Endometrioid tumours were divided into two, *i.e.*, grade 1 or 2, and grade 3, as it has been suggested that some high-grade (grade 3) endometrioid ovarian carcinomas might be the same as high-grade serous carcinomas²⁰ and should be considered with them.⁴

Clear cell tumours were not divided by grade, as all clear cell ovarian cancers are high-grade/grade 3, by definition.^{4,20,21}

Mucinous tumours were split into mucinous borderline tumours (ICD-O codes 8472–8473) and mucinous carcinoma (ICD-O codes 8470–8471 and 8480–8490); the mucinous carcinomas were not further subdivided by tumour grade, as numbers were small.

Statistical analysis

Women were excluded from the analyses if: (i) they had been diagnosed with any invasive cancer other than non-melanoma skin cancer (ICD-10 code C44) prior to recruitment ($n = 66,221$), (ii) they reported at recruitment having had both ovaries removed (bilateral oophorectomy), or if they were unsure whether they had or not ($n = 170,769$), or (iii) they had missing data on tubal ligation ($n = 15,712$). The remaining women ($N = 1,132,914$) contributed person-years from the date of recruitment into the study until the date of registration for ovarian cancer, the date of death, or last date of follow-up (December 31, 2013)—whichever was soonest. Women were censored at diagnosis of any nonovarian cancer. For analyses exploring effects of the timing of tubal ligation, women were excluded if they had missing information on age at tubal ligation ($n = 14,684$). About 1% of participants had been lost to follow-up and such women are censored at the date when they were lost, contributing person-years until then.

Cox (proportional hazards) regression models were used to estimate hazard ratios [referred to as relative risks (RRs)] of developing ovarian cancer by tubal ligation status. Attained age was the underlying time variable. There was no evidence of significant violation of the proportional hazards assumption, as assessed by graphical methods and tests based on Schoenfeld residuals.

All analyses were stratified by geographical region (10 regions corresponding to the areas covered by the cancer

Table 1. Characteristics of the study population at recruitment, and details of follow-up, by tubal ligation status

Characteristics	Tubal ligation		All women
	No	Yes	
Number of women	886,866	246,048	1,132,914
Mean (SD) age at recruitment (years)	56.3 (4.9)	55.3 (4.3)	56.1 (4.8)
Socioeconomic status, lower third, % (n)	30.5 (268,314)	38.0 (92,663)	32.1 (360,977)
Mean age at menarche (SD)	13.0 (1.6)	13.0 (1.6)	13.0 (1.6)
Nulliparous, % (n)	13.3 (117,319)	2.6 (6,352)	10.9 (123,671)
Ever use of oral contraceptive pill, % (n)	57.0 (501,857)	70.3 (171,546)	59.9 (673,403)
Ever use of menopausal hormones, % (n)	44.5 (390,518)	54.9 (133,529)	46.7 (524,047)
Hysterectomy, % (n)	14.2 (126,054)	19.4 (47,579)	15.4 (173,633)
Mean age at natural menopause (SD)	49.3 (4.2)	48.7 (4.5)	49.1 (4.3)
Family history of breast cancer, % (n)	9.8 (82,101)	10.0 (22,781)	9.9 (104,882)
Mean (SD) body mass index (kg/m ²)	26.0 (4.6)	26.6 (4.8)	26.1 (4.7)
Current smoker, % (n)	18.4 (153,773)	26.4 (61,264)	20.1 (215,037)
Strenuous exercise $\geq$ once/week, % (n)	39.8 (340,538)	37.7 (89,266)	39.3 (429,804)
Alcohol intake, $\geq$ 7 units/week, % (n)	23.6 (207,870)	25.5 (62,130)	24.0 (270,000)
Follow-up for incident ovarian cancer			
Woman-years of follow-up for incidence (100,000s)	122.6	33.9	156.5
Mean follow-up time per woman (SD)	13.8 (3.4)	13.8 (3.3)	13.8 (3.4)
Number of incident ovarian cancers	6,693	1,342	8,035

Means and percentages are calculated excluding missing values for the variable of interest.
 n: number of women. SD: standard deviation.

registries), and further adjusted for parity (0, 1, 2, 3+), use of the oral contraceptive pill (never, ever), family history of breast cancer (no, yes), hysterectomy (no, yes), use of menopausal hormones (never, ever), body mass index (BMI) (<25 kg/m², 25–29 kg/m², 30+ kg m⁻²), smoking history (never, past, current) and quintiles of socioeconomic status (based on the Townsend deprivation index).²² All adjustment variables were as reported at recruitment. For adjustment variables, missing values were assigned to a separate category. Exposure information was either missing or reported as unknown for $\leq$ 6% of women for all potential confounders. Sensitivity analyses were conducted excluding all women with missing data on covariates within the model.

Other factors (including alcohol consumption, physical activity and age at menarche) were explored as potential confounders, but were not included in the final model as their inclusion made no appreciable difference to the main estimate of effect.

Tests of heterogeneity in the relationship between tubal ligation and ovarian cancer risk by histotype were performed using a competing risks approach.²³

To explore possible effect modification of the main relationship between tubal ligation and ovarian cancer, multiplicative interaction terms were modelled between tubal ligation and each potential confounder in turn. Likelihood ratio tests were used to compare the models with versus without the interaction terms.

Analyses were also conducted to explore possible effects of age and time with tubal ligation, including age at, time since and calendar year of tubal ligation, and whether the tubal ligation was in the same year as the last birth or afterward. All analyses of the timing of tubal ligation were restricted to parous women, as there were very few nulliparous women who had had a tubal ligation. Where comparisons are made between more than two exposure categories, group-specific confidence intervals for the log risk in each group were calculated, allowing comparisons to be made between any two categories, even if neither is the reference group.²⁴ Conventional 95% confidence intervals (CI) are given in the text.

Analyses were performed in Stata-14.²⁵ Tests of statistical significance were two-sided. Figures were drawn in R using Matthew Arnold and Paul Sherliker's "Jasper" package.

Results

A total of 1,132,914 women, mean age at recruitment 56.1 [standard deviation (SD) 4.8] years, were included in the analyses. At recruitment, 246,048 (22%) reported having had a tubal ligation, at median age 35 [interquartile range (IQR) 31–38], and median year 1978 (IQR 1973–1982). Women who reported a previous tubal ligation were more likely to be parous, to have used the oral contraceptive pill and menopausal hormones, to have had a hysterectomy, to be a current smoker, and to live in areas in the lower third of

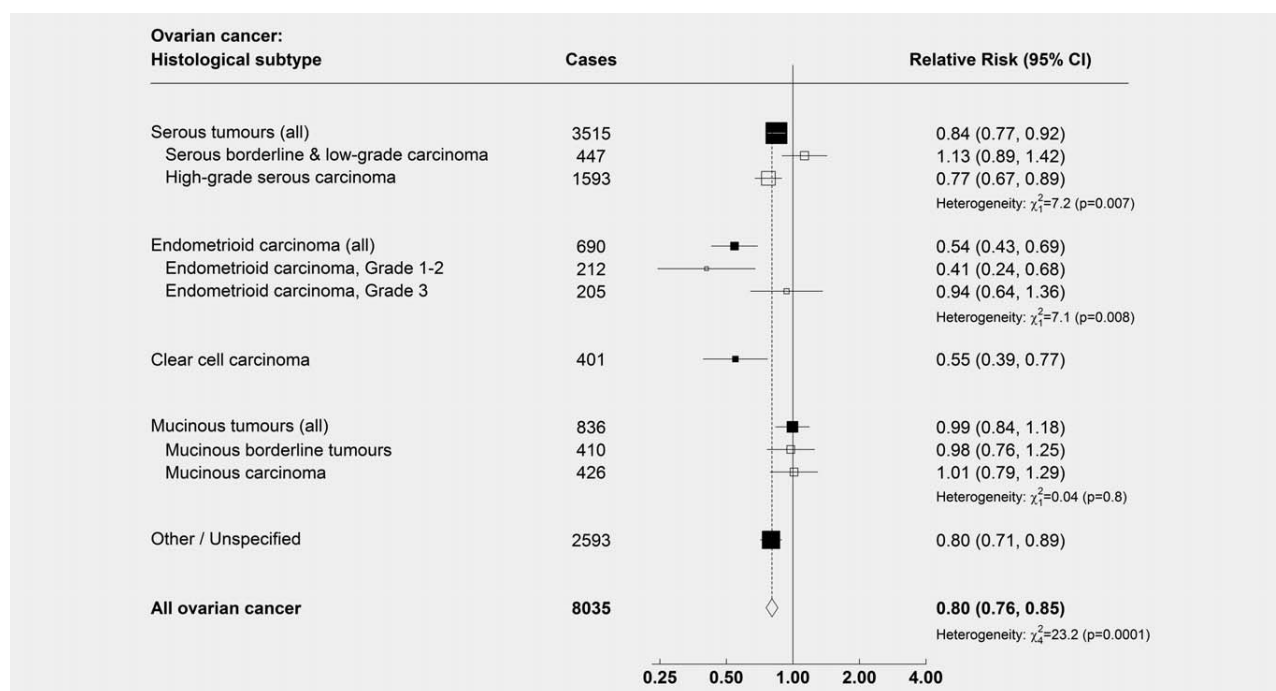


Figure 1. Relative risk of subtypes of ovarian cancer in women with versus without a history of tubal ligation. $N = 1,132,914$. Results show relative risks (hazard ratios) of ovarian cancer amongst women with a history of tubal ligation compared to women without a history of tubal ligation, by histological subtype. Analyses are adjusted for age, region, parity, family history of breast cancer, hysterectomy, use of the oral contraceptive pill and menopausal hormones, body mass index, smoking and socioeconomic status. Serous tumours have been split into low-grade (serous borderline tumours and low-grade serous carcinoma, here defined as grade 1) and high-grade (serous carcinoma of grade ≥ 2). Endometrioid tumours have been split into low-grade (here defined as grade 1–2) and high-grade (grade 3); see main text for discussion of rationale for this. Note: The numbers of grade-specific serous and endometrioid tumours do not sum to the total numbers of serous and endometrioid tumours, as information on tumour grade was missing for 1,475 serous carcinomas and 273 endometrioid carcinomas.

socioeconomic status, compared to women without tubal ligation (Table 1).

The women were followed up for incident ovarian cancer over 15.6 million person-years, with a mean duration of follow-up of 13.8 (SD 3.4) years per woman. During this period, 8,035 incident ovarian cancers were registered, of which 3,515 (44%) were serous; 836 (10%) were mucinous; 690 (9%) were endometrioid; 401 (5%) were clear cell and 2,593 (32%) were of other histological types (mostly unspecified epithelial tumours). The mean age at diagnosis of ovarian cancer was 65.1 years (SD 6.4).

The relative risk of ovarian cancer amongst women with tubal ligation compared to those without was 0.80 (95% CI: 0.76–0.85, $p < 0.001$), after adjustment for age, region, parity, family history of breast cancer, hysterectomy, socioeconomic status, body-mass index, smoking and use of contraceptive or menopausal hormones.

There was strong evidence of heterogeneity by histotype (Fig. 1; heterogeneity, $p = 0.0001$). Tubal ligation was associated with almost a halving in the risk of endometrioid (RR: 0.54, 95% CI: 0.43–0.69) and clear cell tumours (RR: 0.55, 95% CI: 0.39–0.77), and a lesser but still significant reduction in the risk of serous tumours (RR: 0.84, 95% CI: 0.77–0.92),

but there was no significant reduction in the risk of mucinous tumours (RR: 0.99, 95% CI: 0.84–1.18).

When we divided serous tumours into low-grade serous tumours (serous borderline tumours and low-grade serous carcinoma) and high-grade serous carcinoma, risks differed significantly (heterogeneity, $p = 0.007$). There was strong evidence that tubal ligation was associated with a reduction in the risk of high-grade serous carcinoma ($n = 1,593$, RR: 0.77, 95% CI: 0.67–0.89), but not of low-grade serous tumours ($n = 447$, RR: 1.13, 95% CI: 0.89–1.42).

In view of suggestions that some tumours classified as high-grade endometrioid tumours might have been mis-diagnosed high-grade serous tumours,⁴ we conducted an exploratory analysis, splitting endometrioid tumours into low-grade carcinoma, and high-grade carcinoma. There was again heterogeneity in the risks ($p = 0.008$), with tubal ligation associated with a greater reduction in the risk of low-grade ($n = 212$, RR: 0.41, 95% CI: 0.24–0.68), than of high-grade endometrioid carcinoma ($n = 205$, RR: 0.94, 95% CI: 0.64–1.36).

When mucinous tumours were split into borderline and fully malignant subtypes, there was little or no association between tubal ligation and the risk of either mucinous borderline tumours ($n = 410$, RR: 0.98, 95% CI: 0.76–1.25) or

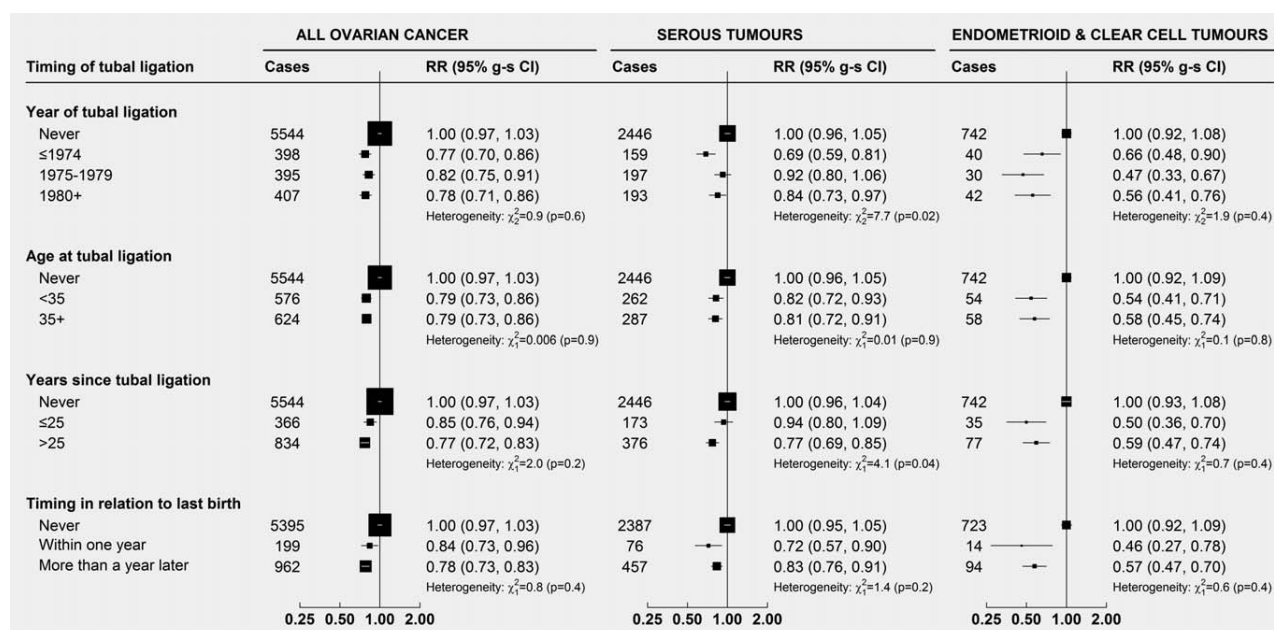


Figure 2. Relative risk of ovarian cancer in relation to the timing of tubal ligation (amongst parous women only). Results show relative risks (hazard ratios) of ovarian cancer by the timing of tubal ligation, with 95% group-specific confidence intervals (95% g-s CI). Each timing analysis was a separate model. For each analysis, we excluded nulliparous women, and sterilized women if their age at tubal ligation was unknown [993,166 women (6,744 cases) included in analyses]. For the analysis of tubal ligation at versus after the last birth, women with an unknown age at last birth were also excluded [967,166 women (6,556 cases) included in analyses]. Analyses are adjusted for age, region, parity, family history of breast cancer, hysterectomy, use of the oral contraceptive pill and menopausal hormones, body mass index, smoking and socioeconomic status. Heterogeneity tests are among women with a tubal ligation only.

fully malignant mucinous carcinoma ($n = 426$, RR: 1.01, 95% CI: 0.79–1.29).

A sensitivity analysis excluding all women with missing values in any of the adjustment variables showed that the association between tubal ligation and ovarian cancer, and the variation by histotype, was not appreciably changed (data not shown).

There was little or no variation in risk either by age at, year of, or years since tubal ligation, or whether tubal ligation was performed in the same year as the last birth or subsequently, for ovarian cancer overall, and for endometrioid and clear cell tumours (Fig. 2). Compared to women without tubal ligation, for serous tumours there was some suggestion of a greater reduction in risk with tubal ligations performed ≤1974 (RR: 0.69, 95% CI: 0.58–0.81), than with those performed in 1975–1979 (RR: 0.92, 95% CI: 0.80–1.07), or those performed in 1980 or later (RR: 0.84, 95% CI: 0.73–0.98), heterogeneity: $p = 0.02$. There was also some suggestion of a greater reduction in risk for women with a time since tubal ligation of >25 years (RR: 0.77, 95% CI: 0.69–0.86) than for ≤25 years since tubal ligation (RR: 0.94, 95% CI: 0.80–1.10), heterogeneity: $p = 0.04$.

There was no significant variation in the association between tubal ligation and ovarian cancer with parity, use of contraceptive or menopausal hormones, hysterectomy, age at natural menopause, family history of breast cancer, smoking, socioeconomic status, frequency of strenuous exercise or

alcohol intake (Fig. 3). There was some weak evidence of heterogeneity by body mass index and age at menarche, although this may well be due to chance, as multiple statistical tests were performed.

Discussion

In a large prospective cohort study with 8,035 incident cases of ovarian cancer, we found strong evidence that the risk of ovarian cancer associated with tubal ligation varied by tumour histotype. We found a modest reduction in risk of the most common histotype, serous tumours, as had been previously reported.^{7–9} However, when we split serous tumours into low-grade serous tumours and high-grade serous carcinoma, we found significant differences, with a 20% reduction in risk of high-grade serous carcinoma, but no reduction in risk of low-grade serous tumours. This is the first study to report a significant difference between the risks of low-grade versus high-grade serous ovarian tumours associated with tubal ligation.

We found about a halving of the risk of endometrioid and of clear cell tumours associated with tubal ligation. Others have reported similar findings for endometrioid tumours.^{7–9} Reported findings for clear cell tumours are mixed, perhaps because it is relatively uncommon. We found little or no association between tubal ligation and the risk of mucinous ovarian cancer, in keeping with some reports,^{7,8} though not others.⁹

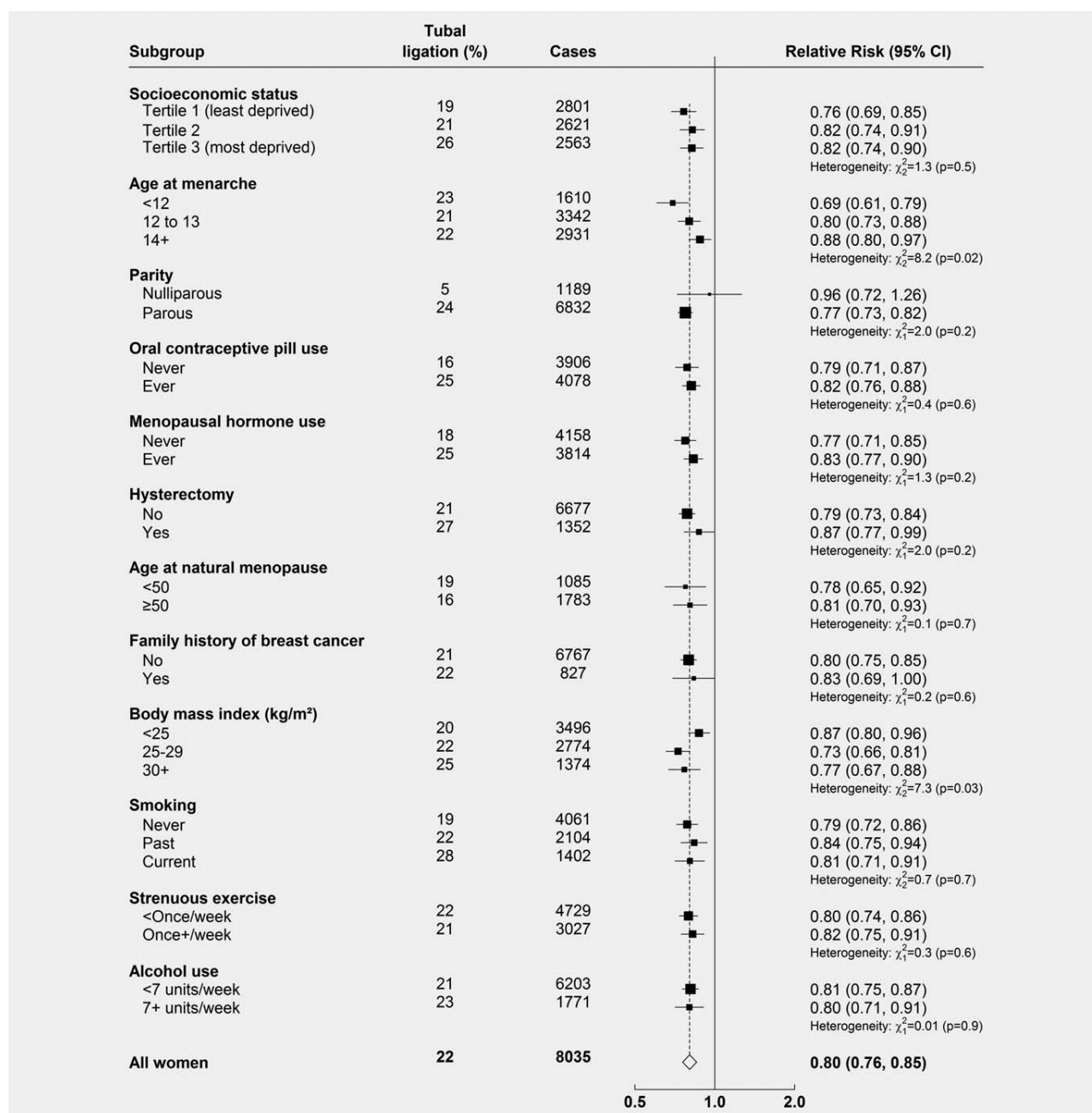


Figure 3. Relative risk of ovarian cancer in women with versus without a history of tubal ligation, by subgroup. Results show the relative risks (hazard ratios) of ovarian cancer in women with versus without tubal ligation, by various other factors. The results for each subgroup are from separate models, restricted to women with no missing information on that variable, and thus total numbers of participants and cases will differ. Analyses are adjusted for age, region, parity, family history of breast cancer, hysterectomy, use of the oral contraceptive pill or menopausal hormones, body mass index, smoking and socioeconomic status, as appropriate.

New hypotheses about the origins of ovarian cancer have provoked a reconsideration of possible mechanisms underlying the reduced risk of ovarian cancer with tubal ligation.^{7,9,26,27} The significant heterogeneity seen here between low-grade versus high-grade serous tumours is consistent with the hypothesis that the two tumour subtypes develop along distinct pathways.

High-grade serous ovarian cancer is hypothesised to arise from precursor lesions within the fallopian tubal epithelium, particularly in the fimbrial end adjacent to the ovary, and subsequently seed to the ovary. There is considerable evidence for this, including several studies showing putative precursor lesions within the fallopian tubes (such as focal areas of epithelium featuring mutations in the tumour-suppressor

gene *TP53* or aberrant expression of its protein p53, and/or areas of dysplasia or intraepithelial serous carcinoma), not only in prophylactic salpingoophorectomy specimens from women with an increased genetic risk of ovarian cancer, and from women with known high-grade serous ovarian cancer,^{28–34} but also in salpingoophorectomy specimens from women not known to be at increased genetic risk, in some cases with accompanying incidental invasive high-grade serous carcinoma.^{35–37} The putative precursor lesions in many cases show molecular changes that are similar or identical to those seen in co-existing high-grade serous ovarian cancer, particularly in terms of *TP53* mutations³³ (mutations in the gene for p53 being characteristic and ubiquitous in high-grade serous ovarian cancer³⁸).

The origins of low-grade serous tumours are considerably more uncertain. In many cases, there appears to be a step-wise progression from ovarian epithelial inclusion cysts, through benign cystadenomas and serous borderline tumours, to low-grade serous carcinoma. Evidence for this includes the frequent co-existence of serous borderline tumours with low-grade serous carcinomas,¹⁸ and similar molecular changes: both serous borderline tumours and low-grade serous carcinomas commonly feature mutations in *BRAF* or *KRAS*,³⁹ but rarely in *TP53*,⁴⁰ in contrast to high-grade serous carcinoma. It has been suggested that some low-grade serous tumours may also originate indirectly from tubal epithelial cells (perhaps from benign or hyperplastic tubal epithelial cells which become trapped in the ovary as epithelial inclusion cysts at the time of ovulation, possibly at a young age),^{41,42} but the evidence for this is much less established than that for the tubal origins of high-grade serous carcinoma.

Surgical ligation of the fallopian tube might reduce the risk of high-grade serous ovarian cancer, either by physically obstructing the passage of tubal cells toward the ovary, or by collateral damage to the local tubal blood supply, with potential subsequent effects on the tubal epithelium. In addition, some forms of tubal sterilization involve removal of some or all of the tube, and/or disruption of the distal end of the tube—which might have been more common in the past, and could possibly explain the significantly greater risk reduction found for tubal ligations done before than after 1974.

The apparent lack of an effect of tubal ligation for low-grade serous tumours is consistent with the hypothesis that they have a different origin to high-grade serous ovarian cancer. However, it is not entirely incompatible with a possible tubal origin for low-grade serous tumours (e.g., if the precursor cells of origin had already been attached to the ovary before the tubal ligation occurred).

Some endometrioid and clear cell tumours are thought to develop from endometriosis,^{3,4} and one might speculate that the almost halving of risk seen with tubal ligation could be due to the blocking of retrograde menstruation through the fallopian tubes to the ovaries and pelvis. It could also reflect the blocking of the transport of carcinogens up the tubes to the ovaries and pelvis. For any mechanism that involved the

blockage of menstruation or the passage of carcinogens up the tube, one might expect the association with tubal ligation to vary by timing (e.g., a greater reduction in risk for tubal ligation performed at a younger age, or longer ago). However, we found no suggestion of any association with the timing of tubal ligation for ovarian cancer overall or for endometrioid and clear cell tumours, and only limited evidence of an association with the year of, and years since, tubal ligation for serous tumours. These findings are in keeping with most, but not all, previous studies.^{7–9}

Our study has several strengths, in addition to the large sample size and number of cases. In particular, the prospective collection of exposure data helped to prevent differential recall of tubal ligation or other factors amongst women with and without cancer. The use of routinely collected national data for follow-up and ascertainment of incident ovarian cancer resulted in few participants being lost to follow-up. The prospective study design also facilitated the involvement of women with more aggressive tumour types, who might not be included in retrospective studies, as there can be a substantial time lag between cancer diagnosis and recruitment to a retrospective study. The high prevalence of tubal ligation (22%) was also an advantage in terms of statistical power.

One potential weakness of our study was that we did not undertake central histopathological review of our cases, for standardization and incorporation of the latest diagnostic criteria. The distribution of histotypes amongst our cases was consistent with those from other population-based studies, with serous tumours accounting for around 65% of those of the four main histotypes, endometrioid tumours for 13%, mucinous 15%, and clear cell 7% (by way of comparison, the corresponding percentages from the EPIC cohort are: 67% serous, 13% endometrioid, 14% mucinous, and 6% clear cell⁴³). This distribution is slightly different to that reported by two hospital-based studies of ovarian cancer cases, in both of which mucinous tumours accounted for only around 3% of cases^{44,45}; this may reflect differences in the study populations (North America vs. Europe), or historical differences in diagnostic criteria (e.g., mucinous tumours in our cohort might include some metastases from the gastrointestinal tract).

About 32% of ovarian cancers in our study were not recorded as being of one of the four main histotypes (serous, endometrioid, clear cell or mucinous). Of these, 87% were epithelial tumours of other, mixed, or unspecified type (mostly unspecified adenocarcinoma, ICD-O code 8140, or unspecified carcinoma, ICD-O code 8010), and 10% were unspecified malignant tumours (ICD-O code 8000). Given that high-grade serous carcinoma is the most common histotype of ovarian cancer, it seems likely that many of the tumours listed as unspecified carcinoma/adenocarcinoma are in fact high-grade serous tumours. This would be supported by our observation that tubal ligation is associated with a similar reduction in risk of other/unspecified ovarian cancer (RR: 0.80, 95% CI: 0.71–0.89) as of high-grade serous carcinoma (RR: 0.77, 95% CI: 0.67–0.89).

Tumour histology may have been classified in somewhat different ways by different pathologists. Any such misclassification would tend to blur differences by tumour histotype, yet distinctly heterogeneous risks were found. This argues strongly for causality, as it confirms that the variation in risk associated with tubal ligation is not just due to confounding, and strengthens hypotheses that different histotypes of ovarian cancer have different causes.

Acknowledgements

The authors thank all the women who participated in the Million Women Study, and the staff from the NHS Breast Screening Centres. They also thank the anonymous reviewers for their very helpful comments and suggestions. Some of these results were presented in preliminary form at the National Cancer Research Institute Cancer Conference, 2014 (Liverpool, UK). This work was supported by Cancer Research UK (CRUK) grant number C38302/A17318, through a CRUK Oxford Centre Clinical Research Training Fellowship for KG.

Million Women Study Collaborators

Million Women Study Co-ordinating Centre staff: Hayley Abbiss, Simon Abbott, Rupert Alison, Naomi Allen, Miranda Armstrong, Krys Baker, Angela Balkwill, Emily Banks, Isobel Barnes, Valerie Beral, Judith Black, Roger Blanks, Kathryn Bradbury, Anna Brown, Benjamin Cairns, Karen Canfell, Dexter Canoy, Andrew Chadwick, Barbara Crossley, Francesca Crowe, Dave Ewart, Sarah Ewart, Lee Fletcher, Sarah Floud, Toral Gathani, Laura Gerrard, Adrian Goodill, Jane Green,

Lynden Guiver, Michal Hozak, Isobel Lingard, Sau Wan Kan, Oksana Kirichek, Nicky Langston, Bette Liu, Kath Moser, Kirstin Pirie, Gillian Reeves, Keith Shaw, Emma Sherman, Helena Strange, Sian Sweetland, Sarah Tipper, Ruth Travis, Lyndsey Trickett, Lucy Wright, Owen Yang, Heather Young.

Million Women Study Advisory Committee: Emily Banks, Valerie Beral, Lucy Carpenter, Carol Dezateux, Jane Green, Julietta Patnick, Richard Peto, Cathie Sudlow.

The NHS Breast Screening Centres which took part in the recruitment of participants were: Avon, Aylesbury, Barnsley, Basingstoke, Bedfordshire and Hertfordshire, Cambridge and Huntingdon, Chelmsford and Colchester, Chester, Cornwall, Crewe, Cumbria, Doncaster, Dorset, East Berkshire, East Cheshire, East Devon, East of Scotland, East Suffolk, East Sussex, Gateshead, Gloucestershire, Great Yarmouth, Hereford and Worcester, Kent, Kings Lynn, Leicestershire, Liverpool, Manchester, Milton Keynes, Newcastle, North Birmingham, North East Scotland, North Lancashire, North Middlesex, North Nottingham, North of Scotland, North Tees, North Yorkshire, Nottingham, Oxford, Portsmouth, Rotherham, Sheffield, Shropshire, Somerset, South Birmingham, South East Scotland, South East Staffordshire, South Derbyshire, South Essex, South Lancashire, South West Scotland, Surrey, Warrington Halton St Helens and Knowsley, Warwickshire Solihull and Coventry, West Berkshire, West Devon, West London, West Suffolk, West Sussex, Wiltshire, Winchester, Wirral, Wycombe.

References

- Forman D, Bray F, Brewster D, et al. eds. Cancer incidence in five continents, vol. 10. Lyon: International Agency for Research on Cancer (IARC Scientific Publication No. 164.ed.), 2014.
- Kurman RJ, Shih IeM. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer—shifting the paradigm. *Hum Pathol* 2011;42: 918–31.
- Prat J. Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features. *Virchows Arch* 2012; 460:237–49.
- Kurman RJ, Carcangiu ML, Herrington CS, et al., eds. WHO classification of tumours of female reproductive organs, 4th edn. Lyon: IARC, 2014.
- Kobel M, Kalloger SE, Boyd N, et al. Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. *PLoS Med* 2008;5:e232.
- McCluggage WG. Morphological subtypes of ovarian carcinoma: a review with emphasis on new developments and pathogenesis. *Pathology* 2011;43:420–32.
- Cibula D, Widschwendter M, Majek O, et al. Tubal ligation and the risk of ovarian cancer: review and meta-analysis. *Hum Reprod Update* 2011;17:55–67.
- Rice MS, Murphy MA, Tworoger SS. Tubal ligation, hysterectomy and ovarian cancer: a meta-analysis. *J Ovarian Res* 2012;5:13.
- Sieh W, Salvador S, McGuire V, et al. Tubal ligation and risk of ovarian cancer subtypes: a pooled analysis of case-control studies. *Int J Epidemiol* 2013;42:579–89.
- The Million Women Study Collaborative Group. The million women study: design and characteristics of the study population. The Million Women Study Collaborative Group. *Breast Cancer Res* 1999;1:73–80.
- Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* 2003;362:419–27.
- World Health Organization. International statistical classification of diseases and related health problems, 10th revision edn. Geneva: World Health Organization, 1992.
- Percy C, Van Holten V, Muir CS. International classification of diseases for oncology: ICD-O, 2nd edn. Geneva: World Health Organization, 1990.
- Fritz A, Percy C, Jack A, et al. International classification of diseases for oncology: ICD-O, 3rd edn. Geneva: World Health Organization, 2000.
- Shepherd JH. Revised FIGO staging for gynaecological cancer. *Br J Obstet Gynaecol* 1989;96:889–92.
- Tavassoli FA, Devilee P, eds. World health organization classification of tumours. Pathology and genetics of tumours of the breast and female genital organs ed. Lyon: IARC Press, 2003.
- Silverberg SG. Histopathologic grading of ovarian carcinoma: a review and proposal. *Int J Gynecol Pathol* 2000;19:7–15.
- Malpica A, Deavers MT, Lu K, et al. Grading ovarian serous carcinoma using a two-tier system. *Am J Surg Pathol* 2004;28:496–504.
- Bodurka DC, Deavers MT, Tian C, et al. Reclassification of serous ovarian carcinoma by a 2-tier system: a gynecologic oncology group study. *Cancer* 2012;118:3087–94.
- Prat J, for the FIGO Committee on Gynecologic Oncology. FIGO guidelines staging classification for cancer of the ovary, fallopian tube, and peritoneum. *Int J Gynaecol Obstet* 2014;124:1–5.
- Wilkinson N, McCluggage WG, on behalf of the Royal College of Pathologists' Working Group on Cancer Services. Datasets for the histopathological reporting of neoplasms of the ovaries and fallopian tubes and primary carcinomas of the peritoneum, 3rd edn. London: Royal College of Pathologists, 2010.
- Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the north ed. London: Croom Helm, 1988.
- Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics* 1995;51:524–32.
- Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med* 1991;10:1025–35.
- StataCorp. Stata Statistical Software: release, 14th edn. College Station, TX: StataCorp LP, 2015.
- Madsen C, Baandrup L, Dehlendorff C, et al. Tubal ligation and salpingectomy and the risk of epithelial ovarian cancer and borderline ovarian tumors: a nationwide case-control study. *Acta Obstet Gynecol Scand* 2015;94:86–94.
- Cibula D, Widschwendter M, Zikan M, et al. Underlying mechanisms of ovarian cancer risk reduction after tubal ligation. *Acta Obstet Gynecol Scand* 2011;90:559–63.

28. Piek JM, van Diest PJ, Zweemer RP, et al. Dysplastic changes in prophylactically removed Fallopian tubes of women predisposed to developing ovarian cancer. *J Pathol* 2001;195:451–6.
29. Colgan TJ, Murphy J, Cole DE, et al. Occult carcinoma in prophylactic oophorectomy specimens: prevalence and association with BRCA germline mutation status. *Am J Surg Pathol* 2001;25:1283–9.
30. Medeiros F, Muto MG, Lee Y, et al. The tubal fimbria is a preferred site for early adenocarcinoma in women with familial ovarian cancer syndrome. *Am J Surg Pathol* 2006;30:230–6.
31. Callahan MJ, Crum CP, Medeiros F, et al. Primary fallopian tube malignancies in BRCA-positive women undergoing surgery for ovarian cancer risk reduction. *J Clin Oncol* 2007;25:3985–90.
32. Folkins AK, Jarboe EA, Saleemuddin A, et al. A candidate precursor to pelvic serous cancer (p53 signature) and its prevalence in ovaries and fallopian tubes from women with BRCA mutations. *Gynecol Oncol* 2008;109:168–73.
33. Kindelberger DW, Lee Y, Miron A, et al. Intraepithelial carcinoma of the fimbria and pelvic serous carcinoma: evidence for a causal relationship. *Am J Surg Pathol* 2007;31:161–9.
34. Salvador S, Rempel A, Soslow RA, et al. Chromosomal instability in fallopian tube precursor lesions of serous carcinoma and frequent monoclonality of synchronous ovarian and fallopian tube mucosal serous carcinoma. *Gynecol Oncol* 2008;110:408–17.
35. Gilks CB, Irving J, Kobel M, et al. Incidental non-uterine high-grade serous carcinomas arise in the fallopian tube in most cases: further evidence for the tubal origin of high-grade serous carcinomas. *Am J Surg Pathol* 2015;39:357–64.
36. Morrison JC, Blanco LZ, Jr., Vang R, et al. Incidental serous tubal intraepithelial carcinoma and early invasive serous carcinoma in the nonprophylactic setting: analysis of a case series. *Am J Surg Pathol* 2015;39:442–53.
37. Rabban JT, Garg K, Crawford B, et al. Early detection of high-grade tubal serous carcinoma in women at low risk for hereditary breast and ovarian cancer syndrome by systematic examination of fallopian tubes incidentally removed during benign surgery. *Am J Surg Pathol* 2014;38:729–42.
38. Ahmed AA, Etemadmoghadam D, Temple J, et al. Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J Pathol* 2010;221:49–56.
39. Singer G, Oldt R, III, Cohen Y, et al. Mutations in BRAF and KRAS characterize the development of low-grade ovarian serous carcinoma. *J Natl Cancer Inst* 2003;95:484–6.
40. Singer G, Stohr R, Cope L, et al. Patterns of p53 mutations separate ovarian serous borderline tumors and low- and high-grade carcinomas and provide support for a new model of ovarian carcinogenesis: a mutational analysis with immunohistochemical correlation. *Am J Surg Pathol* 2005;29:218–24.
41. Li J, Abushahin N, Pang S, et al. Tubal origin of 'ovarian' low-grade serous carcinoma. *Mod Pathol* 2011;24:1488–99.
42. Vang R, Shih Ie M, Kurman RJ. Fallopian tube precursors of ovarian low- and high-grade serous neoplasms. *Histopathology* 2013;62:44–58.
43. Gram IT, Lukanova A, Brill I, et al. Cigarette smoking and risk of histological subtypes of epithelial ovarian cancer in the EPIC cohort study. *Int J Cancer* 2012;130:2204–10.
44. Kobel M, Kalloger SE, Huntsman DG, et al. Differences in tumor type in low-stage versus high-stage ovarian carcinomas. *Int J Gynecol Pathol* 2010;29:203–11.
45. Seidman JD, Horkayne-Szakaly I, Haiba M, et al. The histologic type and stage distribution of ovarian carcinomas of surface epithelial origin. *Int J Gynecol Pathol* 2004;23:41–4.

**C.2 TUBAL LIGATION AND INCIDENCE OF 26 SITE-SPECIFIC CANCERS
IN THE MILLION WOMEN STUDY (BRITISH JOURNAL OF CANCER,
2016)**

I am the first author and local guarantor for this paper, which is based on work in Chapter 6. I conducted the statistical analyses for this project, drafted the paper, and revised it in response to reviewers' comments, with input from co-authors.

Keywords: tubal ligation; sterilisation; cancer risk; ovarian cancer; peritoneal cancer; fallopian tube cancer

Tubal ligation and incidence of 26 site-specific cancers in the Million Women Study

Kezia Gaitskell^{*,1}, Kate Coffey¹, Jane Green¹, Kirstin Pirie¹, Gillian K Reeves¹, Ahmed A Ahmed², Isobel Barnes¹ and Valerie Beral¹ on behalf of the Million Women Study collaborators³

¹Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Richard Doll Building, Roosevelt Drive, Oxford OX3 7LF, UK and ²Nuffield Department of Obstetrics and Gynaecology and The Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK

Background: Tubal ligation is known to be associated with a reduction in ovarian cancer risk. Associations with breast, endometrial and cervical cancers have been suggested. We investigated associations for 26 site-specific cancers in a large UK cohort.

Methods: Study participants completed a questionnaire on reproductive and lifestyle factors in 1996–2001, and were followed for cancer and death via national registries. Using Cox regression models, we estimated adjusted relative risks (RRs) for 26 site-specific cancers among women with vs without tubal ligation.

Results: In 1 278 783 women without previous cancer, 167 430 incident cancers accrued during 13.8 years' follow-up. Significantly reduced risks were found in women with tubal ligation for cancers of the ovary (RR = 0.80, 95% CI: 0.76–0.85; $P < 0.001$; $n = 8035$), peritoneum (RR = 0.81, 0.66–0.98; $P = 0.03$; $n = 730$), and fallopian tube (RR = 0.60, 0.37–0.96; $P = 0.04$; $n = 168$). No significant associations were found for endometrial, breast, or cervical cancers.

Conclusions: The reduced risks of ovarian, peritoneal and fallopian tube cancers are consistent with hypotheses of a common origin for many tumours at these sites, and with the suggestion that tubal ligation blocks cells, carcinogens or other agents from reaching the ovary, fallopian tubes and peritoneal cavity.

Tubal ligation (female sterilisation) is known to be associated with a reduced risk of ovarian cancer (Sieh *et al*, 2013). Reports of associations with other cancers have varied. In particular, several investigators have described associations with breast (Irwin *et al*, 1988; Kreiger *et al*, 1999), endometrial (Kjaer *et al*, 2004), cervical (Mathews *et al*, 2012) and anal cancers (Coffey *et al*, 2015). We report here the associations between tubal ligation and risk of 26 site-specific cancers in a large cohort of UK women.

MATERIALS AND METHODS

Study design, data collection and follow-up. The Million Women Study is a prospective study of 1.3 million UK women, recruited in 1996–2001. Participants completed a questionnaire at recruitment, on socio-demographic, reproductive and lifestyle

factors. Information on tubal ligation came from responses to the question 'Have you been sterilised (had your tubes tied)?' on the recruitment questionnaire. The study design and methods have been described previously (The Million Women Study Collaborative Group, 1999; Million Women Study Collaborators, 2003) and questionnaires can be viewed online at <http://www.millionwomenstudy.org>.

Follow-up for outcomes was based on routine registers: participants have been flagged on the NHS Central Register for cancer registrations and deaths. Information on cancer provided to investigators includes the date of registration and cancer site (coded using the 10th revision of the International Classification of Diseases, ICD-10) (World Health Organization, 1992). All participants gave written consent to follow-up at recruitment. Ethical approval was granted by the Oxford and Anglia Multi-Centre Research Ethics Committee (MREC 97/01).

*Correspondence: Dr K Gaitskell; E-mail: kezia.gaitskell@ceu.ox.ac.uk

³Members of the Million Women Study collaborators are listed before the references.

Received 6 January 2016; revised 19 February 2016; accepted 3 March 2016

© 2016 Cancer Research UK. All rights reserved 0007–0920/16

Statistical analysis. Women were excluded from the analyses if they had invasive cancer other than non-melanoma skin cancer (ICD-10 code C44) before recruitment ($n = 66\,221$ (5%)), or missing data on tubal ligation ($n = 40\,612$ (3%)). The remaining women ($n = 1\,278\,783$) contributed person-years from the date of recruitment until the earliest of the date of registration of any cancer (other than non-melanoma skin cancer), the date of death, or last date of follow-up (31st December 2013). About 1% of participants were lost to follow-up and contributed person-years until the date they were lost.

Cox (proportional hazards) regression models were used to estimate hazard ratios (referred to as relative risks (RRs)) of developing a given cancer by tubal ligation status. Attained age was the underlying time variable. Analyses were run separately for each of 26 cancer types or sites, defined by ICD-10 code, and for all cancer (any invasive cancer other than non-melanoma skin cancer, ICD-10 C44). Analyses were limited to cancer types for which 500 or more cases with an identified primary site had accrued. In addition, we conducted an analysis for fallopian tube cancers (ICD-10 C57.0), despite having fewer than 500 cases, as there is a reasonable prior hypothesis that surgery to the fallopian tube might affect the incidence of cancer at this site.

All analyses were stratified by geographical region (10 regions corresponding to the areas covered by the cancer registries) and quintiles of socioeconomic status (based on the Townsend deprivation index) (Townsend *et al*, 1988), and additionally adjusted for parity (0, 1, 2, 3+), age at first birth (<20, 20–24, 25–29, 30+), smoking (never, past, current <10 cigarettes/day, current 10–19 cigarettes/day, current 20+ cigarettes/day), average alcohol intake per week (none, ≤ 7 units, > 7 & ≤ 14 units, > 14 units), frequency of strenuous exercise (<once/week, $\geq$ once/week), use of menopausal hormones (never, ever), use of the oral contraceptive pill (never, ever), hysterectomy (no, yes), and body mass index (<25 kg m⁻², 25–29 kg m⁻², 30+ kg m⁻²).

All adjustment variables were as reported at recruitment. For adjustment and stratification variables, missing values were assigned to a separate category. Exposure information was either missing or reported as unknown for $\leq 6\%$ of women for all potential confounders. Analyses of ovarian cancer risk were restricted to women without bilateral oophorectomy; those of endometrial and cervical cancer were restricted to women without hysterectomy; those of fallopian tube cancer were restricted to women without either bilateral oophorectomy or hysterectomy (as the fallopian tubes might potentially be removed as part of either surgery). Analyses were performed in Stata-14 (StataCorp, 2015). Figures were drawn in R (R Development Core Team, 2015) using Matthew Arnold’s ‘jasper’ package (Arnold, 2015).

RESULTS

The study population included 1 278 783 women without prior cancer, average age 56.1 (s.d. 4.9) at recruitment, 23% of whom reported a tubal ligation, at an average age of 34.5 (s.d. 5.3). Table 1 shows baseline characteristics of the cohort. A total of 167 430 cancers accrued, after a mean follow-up of 13.8 years (s.d. 3.4). The average age at cancer diagnosis was 65.6 years (s.d. 6.5). There was no association between tubal ligation and risk of all cancers combined (RR = 1.00, 95% confidence interval (CI): 0.98–1.01).

Figure 1 displays adjusted RRs and 95% CIs of 26 site- or type-specific cancers among women with vs without tubal ligation. There was a 20% reduction for ovarian cancer risk (RR = 0.80, 95% CI: 0.76–0.85; $P < 0.001$) and a reduction of similar magnitude for peritoneal cancers (RR = 0.81, 95% CI: 0.66–0.98; $P = 0.03$) (ICD-10 code C48, which includes both peritoneal and retroperitoneal cancers). We found a large reduction in risk of cancers of the fallopian tube, but confidence intervals were wide (RR: 0.60, 95% CI: 0.37–0.96, $P = 0.04$).

Table 1. Characteristics of the study population at recruitment, and details of follow-up, by tubal ligation status		
	Tubal ligation	
	No	Yes
Characteristics		
Number of women	984 059	294 724
Mean (s.d.) age at recruitment (years)	56.3 (5.0)	55.4 (4.3)
Socioeconomic status, lowest quintile, % (n)	18.1 (176 985)	25.1 (73 503)
Nulliparous, % (n)	13.3 (130 818)	2.6 (7558)
Ever use of oral contraceptive pill, % (n)	56.5 (551 799)	69.5 (202 511)
Ever use of menopausal hormones, % (n)	47.5 (463 250)	58.2 (169 199)
Hysterectomy, % (n)	21.9 (215 280)	29.1 (84 514)
Mean (s.d.) body mass index (kg m ⁻²)	26.1 (4.6)	26.7 (4.9)
Current smoker, % (n)	18.5 (171 593)	26.9 (74 647)
Strenuous exercise $\geq$ once/week, % (n)	39.5 (375 006)	37.1 (105 032)
Alcohol intake, ≥ 7 units/week, % (n)	23.3 (227 830)	24.8 (72 471)
Follow-up for cancer incidence		
Woman-years of follow-up (100 000 s)	136.0	40.5
Mean (s.d.) follow-up time per woman	13.8 (3.4)	13.8 (3.3)
Number of incident cancers	129 531	37 899
Abbreviations: n = Number of women; s.d. = standard deviation. Notes: Means and percentages are calculated excluding missing values for the variable of interest.		

We also found increased risk of anal cancer (RR = 1.34, 95% CI: 1.11–1.63; $P = 0.002$), as previously reported (Coffey *et al*, 2015).

The small increase in risk of lung cancer (RR = 1.09, 95% CI: 1.05–1.13) may well be due to residual confounding from smoking — as women with tubal ligation were much more likely to be current smokers, and to smoke more cigarettes per day, than women without tubal ligation. There was no association between tubal ligation and lung cancer risk among never-smokers (RR = 1.04; 95% CI: 0.91–1.19).

There were no significant associations between tubal ligation and risk of cancers at the other sites, including cancers of the endometrium, cervix, breast and colorectum.

DISCUSSION

To our knowledge, this is the largest and most comprehensive analysis so far of the association between tubal ligation and subsequent risk of cancer. We found that tubal ligation was significantly associated with reduced risks of ovarian, peritoneal, and fallopian tube cancers, and an increased risk of anal cancer, but was not associated with breast, endometrial or cervical cancer, as had been suggested by others.

The reduction in ovarian cancer risk has been reported by others (Cibula *et al*, 2011; Rice *et al*, 2012; Sieh *et al*, 2013). We previously reported significant heterogeneity by histological type of ovarian cancer, finding that tubal ligation was associated with a modest reduction of serous tumours overall (RR = 0.84, 95% CI: 0.77–0.92), and about a 20% reduction in risk of high-grade serous carcinoma, and almost a halving of risk of endometrioid and clear cell ovarian tumours (Gaitskell *et al*, 2016). These findings are consistent with hypotheses regarding possible different origins of the ovarian cancer histotypes, and that many cases of high-grade serous ovarian cancer may arise from precursor lesions within the fallopian tubes, while some endometrioid and clear cell tumours may develop from endometriosis (Kurman and Shih, 2011; Prat, 2012). Although the precise mechanism by which tubal ligation reduces the risk of ovarian cancer is unclear, part of the explanation could be that it acts as a physical barrier to pro-cancerous substances passing through the fallopian tubes to the ovaries (whether endometriosis, fallopian tube epithelial cells, infectious agents, or exogenous carcinogens).

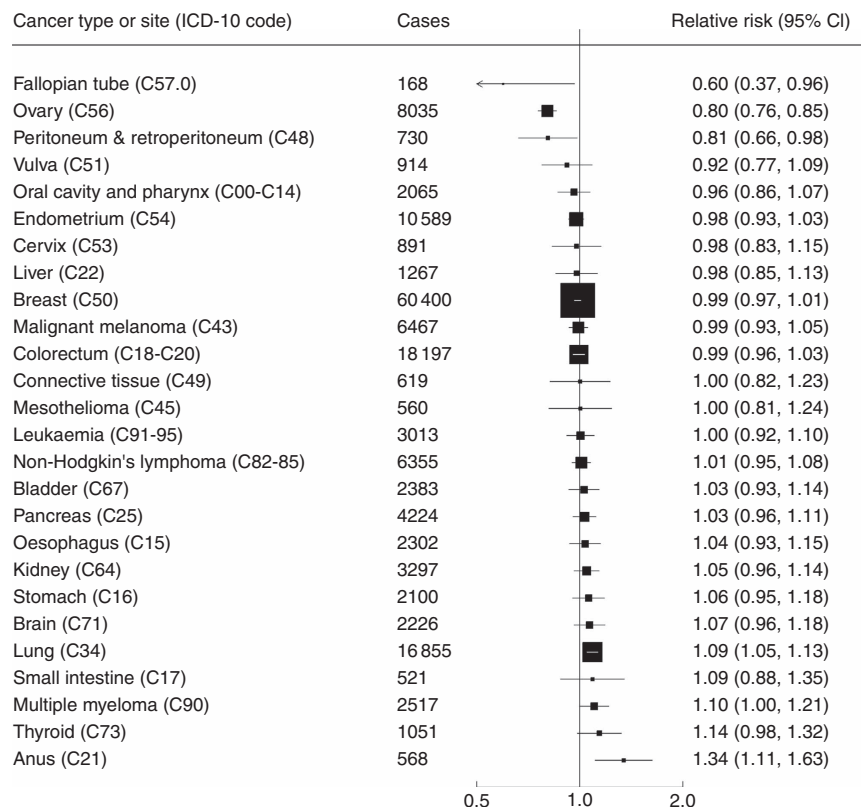


Figure 1. Relative risk of cancer incidence by site or type among women with vs without tubal ligation. Results are adjusted for age, region, socioeconomic status, parity, age at first birth, hysterectomy, smoking, alcohol intake, physical activity, body mass index, and use of the oral contraceptive pill or menopausal hormones. The analysis of fallopian tube cancer is restricted to women without bilateral oophorectomy or hysterectomy. The analysis of ovarian cancer is restricted to women without bilateral oophorectomy. Analyses of endometrial and cervical cancer are restricted to women without hysterectomy.

Few other studies have reported on the association between tubal ligation and risk of peritoneal cancer, and they have tended to be limited by small numbers of cases. Our study included 730 cases of peritoneal cancer; two retrospective studies have reported their findings—one study including 62 cases found no association (Grant *et al*, 2010); and another including 22 cases reported a reduced risk (Purdie *et al*, 1995; Green *et al*, 1997).

We believe that this is the first study to report that tubal ligation is associated with a significant reduction in risk of fallopian tube cancer (Riska *et al*, 2007; Vicus *et al*, 2010).

Our finding that tubal ligation is associated with a reduction in risk of ovarian, peritoneal and fallopian tube cancers perhaps reflects their histological and clinical similarity, and possible similar site of origin (Kurman *et al*, 2014). The majority of the peritoneal and fallopian tube cancers with specified histology were serous carcinomas, which is also the most common histotype of ovarian cancer (Gaitskill *et al*, 2016). There is increasing evidence from histological, molecular and genetics studies that many high-grade serous carcinomas found in the ovary or peritoneum may have originated from precursor lesions in the fallopian tube (reviewed in (Nik *et al*, 2014; Perets and Drapkin, 2016)).

As the majority of the putative precursor lesions for high-grade serous cancers are found within the fimbrial end of the fallopian tube (Przybycin *et al*, 2010; Gilks *et al*, 2015), adjacent to the ovary, it is perhaps surprising that tubal ligation (which normally involves surgery to the mid-portion of the tube) should reduce the risk of such tumours. It may be that, historically, some surgical sterilisation procedures involved cutting and removing a portion of the fallopian tube, rather than simply clipping the tube (as is now more common). Alternatively, it may be that surgery causes local vascular changes with consequent effects on the tubal

epithelium. One of the limitations of our study is that we do not have information on the type of surgical procedure performed, and thus are unable to explore possible differences in association between different surgical techniques.

For endometrial cancer, a reduced risk associated with tubal ligation was reported in a Danish cohort (Kjaer *et al*, 2004), but we and others have not replicated this association (Castellsague *et al*, 1996; Lacey *et al*, 2000). One group of investigators reported that tubal ligation was associated with an increased risk of cervical cancer (Mathews *et al*, 2012), but again, neither other findings (Li and Thomas, 2000) nor ours replicated this.

Reports of the relationship between tubal ligation and breast cancer risk have been varied, including suggestions of either a reduced (Kreiger *et al*, 1999; Calle *et al*, 2001) or increased risk (Irwin *et al*, 1988), although the majority of studies reported no significant association (Brinton *et al*, 2000; Eliassen *et al*, 2006; Iversen *et al*, 2007; Dorjgochoo *et al*, 2009; Press *et al*, 2011; Gaudet *et al*, 2013; Nichols *et al*, 2013), consistent with our findings.

We have previously reported in detail on the risk factors for anal cancer in the Million Women Study (Coffey *et al*, 2015); factors associated with an increased risk of anal cancer included tubal ligation, nulliparity, smoking, past use of the oral contraceptive pill, and not living with a partner. The strongest association was with a history of cervical intraepithelial neoplasia grade 3 (RR = 4.03, 95% CI: 2.59–6.28), reflecting the importance of high-risk strains of human papilloma virus as causative agents of both cervical (Walboomers *et al*, 1999) and anal (International Agency for Research on Cancer, 2009) cancers. It is unclear why tubal ligation should be associated with an increased risk of anal cancer, and it may be that this apparent association reflects confounding by sexual behaviour and exposure to human papilloma virus. For

example, as tubal ligation is usually performed for permanent contraception, it is likely to be more common among women who are sexually active, and who thus may be more likely to be exposed to sexually-transmitted infections, such as human papilloma virus, and the consequent increased risk of anal cancer.

There have been reports on possible relationships between tubal ligation and other cancers, including possible reductions in risk of colorectal (Cape and Kreiger, 1999; Rosenblatt *et al*, 2004) and stomach cancer (Dorjgochoo *et al*, 2009), and increased risks of thyroid cancer (Braganza *et al*, 2014) and lymphatic and haematopoietic malignancies (Kjaer *et al*, 2004). We did not replicate any of these findings.

The large size of the cohort, the individual information on possible confounding factors, and the complete and long follow-up, provided reliable estimates of risks associated with tubal ligation for 26 specific cancer sites, even relatively uncommon ones. We found no association between tubal ligation and the risk of cancers of the endometrium, breast, or cervix. By contrast, tubal ligation is associated with a clear reduction in risk of ovarian cancer, a reduction of similar magnitude of peritoneal cancer, and a reduction of fallopian tube cancer. That tubal ligation is associated with a reduced risk of cancers of the ovary, peritoneum and fallopian tube, but not of other hormonally-related cancers, is consistent with the hypothesis that many of the cancers at these sites have a shared origin in the fallopian tube, and that tubal ligation reduces cancer risk by acting as a barrier to cells, carcinogens or other agents reaching the ovary and peritoneal cavity, rather than by affecting hormone levels.

ACKNOWLEDGEMENTS

The authors thank all the women who participated in the Million Women Study, and the staff from the NHS Breast Screening Centres. Authors gratefully acknowledge the past and present members of Professor Ahmed's research group (especially Dr Yiyan Zheng, Dr Karin Hellner, Dr Sandra Herrero Gonzalez, David Mannion, Dr Fabrizio Miranda, Mr Pubudu Pathiraja, and Charlotte Taylor), whose questions and discussions initially provoked these analyses. Authors also thank the reviewers for their helpful comments and suggestions—particularly Dr Steven Narod for proposing that we conduct analyses on fallopian tube cancer. The Million Women Study is funded by Cancer Research UK (Grant No. C570/A16491) and the UK Medical Research Council (Grant No. MR/K02700X/1). This work was supported by Cancer Research UK (CRUK) Grant Number C38302/A17318, through a CRUK Oxford Centre Clinical Research Training Fellowship for KG.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

MILLION WOMEN STUDY COLLABORATORS

Million Women Study Co-ordinating Centre Staff

Hayley Abbiss, Simon Abbott, Rupert Alison, Naomi Allen, Miranda Armstrong, Krys Baker, Angela Balkwill, Emily Banks, Isobel Barnes, Valerie Beral, Judith Black, Roger Blanks, Kathryn Bradbury, Anna Brown, Benjamin Cairns, Karen Canfell, Dexter Canoy, Andrew Chadwick, Barbara Crossley, Francesca Crowe, Dave Ewart, Sarah Ewart, Lee Fletcher, Sarah Floud, Toral Gathani, Laura Gerrard, Adrian Goodill, Jane Green, Lynden Guiver, Michal Hozak, Isobel Lingard, Sau Wan Kan, Oksana Kirichek, Nicky Langston, Bette Liu, Kath Moser, Kirstin Pirie, Gillian Reeves, Keith Shaw, Emma

Sherman, Helena Strange, Sian Sweetland, Sarah Tipper, Ruth Travis, Lyndsey Trickett, Lucy Wright, Owen Yang, Heather Young.

Million Women Study Advisory Committee

Emily Banks, Valerie Beral, Lucy Carpenter, Carol Dezateux, Jane Green, Julietta Patnick, Richard Peto, Cathie Sudlow.

REFERENCES

- Arnold M (2015) Jasper: Jasper makes plots. R package version 2-192.
- Braganza MZ, de Gonzalez AB, Schonfeld SJ, Wentzensen N, Brenner AV, Kitahara CM (2014) Benign breast and gynecologic conditions, reproductive and hormonal factors, and risk of thyroid cancer. *Cancer Prev Res (Phila)* 7(4): 418–425.
- Brinton LA, Gammon MD, Coates RJ, Hoover RN (2000) Tubal ligation and risk of breast cancer. *Br J Cancer* 82(9): 1600–1604.
- Calle EE, Rodriguez C, Walker KA, Wingo PA, Petrelli JM, Thun MJ (2001) Tubal sterilization and risk of breast cancer mortality in US women. *Cancer Causes Control* 12(2): 127–135.
- Cape DB, Kreiger N (1999) Gynaecological surgical procedures and risk of colorectal cancer in women. *Eur J Cancer Prev* 8(6): 495–500.
- Castellsague X, Thompson WD, Dubrow R (1996) Tubal sterilization and the risk of endometrial cancer. *Int J Cancer* 65(5): 607–612.
- Cibula D, Widschwendter M, Majek O, Dusek L (2011) Tubal ligation and the risk of ovarian cancer: review and meta-analysis. *Hum Reprod Update* 17(1): 55–67.
- Coffey K, Beral V, Green J, Reeves G, Barnes I (2015) Lifestyle and reproductive risk factors associated with anal cancer in women aged over 50 years. *Br J Cancer* 112(9): 1568–1574.
- Dorjgochoo T, Shu XO, Li HL, Qian HZ, Yang G, Cai H, Gao YT, Zheng W (2009) Use of oral contraceptives, intrauterine devices and tubal sterilization and cancer risk in a large prospective study, from 1996 to 2006. *Int J Cancer* 124(10): 2442–2449.
- Eliassen AH, Colditz GA, Rosner B, Hankinson SE (2006) Tubal sterilization in relation to breast cancer risk. *Int J Cancer* 118(8): 2026–2030.
- Gaitskell K, Green J, Pirie K, Reeves G, Beral V. on behalf of the Million Women Study Collaborators (2016) Tubal ligation and ovarian cancer risk in a large cohort: Substantial variation by histological type. *Int J Cancer* 138(5): 1076–1084.
- Gaudet MM, Patel AV, Sun J, Teras LR, Gapstur SM (2013) Tubal sterilization and breast cancer incidence: results from the cancer prevention study II nutrition cohort and meta-analysis. *Am J Epidemiol* 177(6): 492–499.
- Gilks CB, Irving J, Kobel M, Lee C, Singh N, Wilkinson N, McCluggage WG (2015) Incidental nonuterine high-grade serous carcinomas arise in the fallopian tube in most cases: further evidence for the tubal origin of high-grade serous carcinomas. *Am J Surg Pathol* 39(3): 357–364.
- Grant DJ, Moorman PG, Akushevich L, Palmieri RT, Bentley RC, Schildkraut JM (2010) Primary peritoneal and ovarian cancers: an epidemiological comparative analysis. *Cancer Causes Control* 21(7): 991–998.
- Green A, Purdie D, Bain C, Siskind V, Russell P, Quinn M, Ward B (1997) Tubal sterilisation, hysterectomy and decreased risk of ovarian cancer. Survey of Women's Health Study Group. *Int J Cancer* 71(6): 948–951.
- International Agency for Research on Cancer (2009) *A review of human carcinogens. Part B: Biological agents: Human Papillomaviruses* Vol. 100B. IARC Press: Lyon, France.
- Irwin KL, Lee NC, Peterson HB, Rubin GL, Wingo PA, Mandel MG (1988) Hysterectomy, tubal sterilization, and the risk of breast cancer. *Am J Epidemiol* 127(6): 1192–1201.
- Iversen L, Hannaford PC, Elliott AM (2007) Tubal sterilization, all-cause death, and cancer among women in the United Kingdom: evidence from the Royal College of General Practitioners' Oral Contraception Study. *Am J Obstet Gynecol* 196(5): 447 e1–447 e8.
- Kjaer SK, Møller M, Brinton LA, Johansen C, Gridley G, Olsen JH (2004) Tubal sterilization and risk of ovarian, endometrial and cervical cancer. A Danish population-based follow-up study of more than 65 000 sterilized women. *Int J Epidemiol* 33(3): 596–602.
- Kreiger N, Sloan M, Cotterchio M, Kirsh V (1999) The risk of breast cancer following reproductive surgery. *Eur J Cancer* 35(1): 97–101.
- Kurman RJ, Carcangiu ML, Herrington CS, Young RH (Eds) (2014) *WHO Classification of Tumours of Female Reproductive Organs*. 4th edn. IARC: Lyon.

- Kurman RJ, Shih IeM (2011) Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer—shifting the paradigm. *Hum Pathol* **42**(7): 918–931.
- Lacey Jr JV, Brinton LA, Mortel R, Berman ML, Wilbanks GD, Twigg LB, Barrett RJ (2000) Tubal sterilization and risk of cancer of the endometrium. *Gynecol Oncol* **79**(3): 482–484.
- Li H, Thomas DB (2000) Tubal ligation and risk of cervical cancer. The World Health Organization Collaborative Study of Neoplasia and Steroid Contraceptives. *Contraception* **61**(5): 323–328.
- Mathews CA, Stoner JA, Wentzensen N, Moxley KM, Tenney ME, Tuller ER, Myers T, Landrum LM, Lanneau G, Zuna RE, Gold MA, Wang SS, Walker JL (2012) Tubal ligation frequency in Oklahoma women with cervical cancer. *Gynecol Oncol* **127**(2): 278–282.
- Million Women Study Collaborators (2003) Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* **362**(9382): 419–427.
- Nichols HB, Baird DD, DeRoo LA, Kissling GE, Sandler DP (2013) Tubal ligation in relation to menopausal symptoms and breast cancer risk. *Br J Cancer* **109**(5): 1291–1295.
- Nik NN, Vang R, Shih IeM, Kurman RJ (2014) Origin and pathogenesis of pelvic (ovarian, tubal, and primary peritoneal) serous carcinoma. *Annu Rev Pathol* **9**: 27–45.
- Perets R, Drapkin R (2016) It's totally tubular...riding the new wave of ovarian cancer research. *Cancer Res* **76**(1): 10–17.
- Prat J (2012) Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features. *Virchows Arch* **460**(3): 237–249.
- Press DJ, Sullivan-Halley J, Ursin G, Deapen D, McDonald JA, Strom BL, Norman SA, Simon MS, Marchbanks PA, Folger SG, Liff JM, Burkman RT, Malone KE, Weiss LK, Spirtas R, Bernstein L (2011) Breast cancer risk and ovariectomy, hysterectomy, and tubal sterilization in the women's contraceptive and reproductive experiences study. *Am J Epidemiol* **173**(1): 38–47.
- Przybycin CG, Kurman RJ, Ronnett BM, Shih Ie M, Vang R (2010) Are all pelvic (nonuterine) serous carcinomas of tubal origin? *Am J Surg Pathol* **34**(10): 1407–1416.
- Purdie D, Green A, Bain C, Siskind V, Ward B, Hacker N, Quinn M, Wright G, Russell P, Wright G, Susil B (1995) Reproductive and other factors and risk of epithelial ovarian cancer: an Australian case-control study. Survey of Women's Health Study Group. *Int J Cancer* **62**(6): 678–684.
- R Development Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria.
- Rice MS, Murphy MA, Tworoger SS (2012) Tubal ligation, hysterectomy and ovarian cancer: A meta-analysis. *J Ovarian Res* **5**(1): 13.
- Riska A, Sund R, Pukkala E, Gissler M, Leminen A (2007) Parity, tubal sterilization, hysterectomy and risk of primary fallopian tube carcinoma in Finland, 1975–2004. *Int J Cancer* **120**(6): 1351–1354.
- Rosenblatt KA, Gao DL, Ray RM, Nelson ZC, Thomas DB (2004) Contraceptive methods and induced abortions and their association with the risk of colon cancer in Shanghai, China. *Eur J Cancer* **40**(4): 590–593.
- Sieh W, Salvador S, McGuire V, Weber RP, Terry KL, Rossing MA, Risch H, Wu AH, Webb PM, Moysich K, Doherty JA, Felberg A, Miller D, Jordan SJ, Goodman MT, Lurie G, Chang-Claude J, Rudolph A, Kjaer SK, Jensen A, Hogdall E, Bandera EV, Olson SH, King MG, Rodriguez-Rodriguez L, Kiemeny LA, Marees T, Massuger LF, van Altena AM, Ness RB, Cramer DW, Pike MC, Pearce CL, Berchuck A, Schildkraut JM, Whittemore AS (2013) Tubal ligation and risk of ovarian cancer subtypes: a pooled analysis of case-control studies. *Int J Epidemiol* **42**(2): 579–589.
- StataCorp (2015) *Stata Statistical Software: Release 14*. StataCorp LP: College Station, TX.
- The Million Women Study Collaborative Group (1999) The Million Women Study: design and characteristics of the study population. *Breast Cancer Res* **1**(1): 73–80.
- Townsend P, Phillimore P, Beattie A (1988) *Health and Deprivation: Inequality and the North*. Croom Helm: London.
- Vicus D, Finch A, Rosen B, Fan I, Bradley L, Cass I, Sun P, Karlan B, McLaughlin J, Narod SA. Hereditary Ovarian Cancer Clinical Study Group (2010) Risk factors for carcinoma of the fallopian tube in women with and without a germline BRCA mutation. *Gynecol Oncol* **118**(2): 155–159.
- Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N (1999) Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* **189**(1): 12–19.
- World Health Organization (1992) *International Statistical Classification of Diseases and Related Health Problems*. 10th revision edn. World Health Organization: Geneva.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

APPENDIX D

PAPERS PUBLISHED DURING DPHIL AS PART OF COLLABORATIONS WITH OTHERS

D.1 MENOPAUSAL HORMONE USE AND OVARIAN CANCER RISK: INDIVIDUAL PARTICIPANT META-ANALYSIS OF 52 EPIDEMIOLOGICAL STUDIES (LANCET, 2015)

This paper resulted from work by the Collaborative Group on Epidemiological Studies of Ovarian Cancer, and I am listed on the writing committee. I contributed to the histological classification, and I read and commented on drafts of the paper.

Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies



Collaborative Group on Epidemiological Studies of Ovarian Cancer*



Summary

Background Half the epidemiological studies with information about menopausal hormone therapy and ovarian cancer risk remain unpublished, and some retrospective studies could have been biased by selective participation or recall. We aimed to assess with minimal bias the effects of hormone therapy on ovarian cancer risk.

Methods Individual participant datasets from 52 epidemiological studies were analysed centrally. The principal analyses involved the prospective studies (with last hormone therapy use extrapolated forwards for up to 4 years). Sensitivity analyses included the retrospective studies. Adjusted Poisson regressions yielded relative risks (RRs) versus never-use.

Findings During prospective follow-up, 12 110 postmenopausal women, 55% (6601) of whom had used hormone therapy, developed ovarian cancer. Among women last recorded as current users, risk was increased even with <5 years of use (RR 1.43, 95% CI 1.31–1.56; $p < 0.0001$). Combining current-or-recent use (any duration, but stopped <5 years before diagnosis) resulted in an RR of 1.37 (95% CI 1.29–1.46; $p < 0.0001$); this risk was similar in European and American prospective studies and for oestrogen-only and oestrogen-progestagen preparations, but differed across the four main tumour types (heterogeneity $p < 0.0001$), being definitely increased only for the two most common types, serous (RR 1.53, 95% CI 1.40–1.66; $p < 0.0001$) and endometrioid (1.42, 1.20–1.67; $p < 0.0001$). Risk declined the longer ago use had ceased, although about 10 years after stopping long-duration hormone therapy use there was still an excess of serous or endometrioid tumours (RR 1.25, 95% CI 1.07–1.46, $p = 0.005$).

Interpretation The increased risk may well be largely or wholly causal; if it is, women who use hormone therapy for 5 years from around age 50 years have about one extra ovarian cancer per 1000 users and, if its prognosis is typical, about one extra ovarian cancer death per 1700 users.

Funding Medical Research Council, Cancer Research UK.

Copyright © Collaborative Group on Epidemiological Studies of Ovarian Cancer. Open Access article distributed under the terms of CC BY.

Introduction

Use of hormone therapy for the menopause increased rapidly during the 1990s, then halved abruptly in the early 2000s after publication of the Women's Health Initiative, a large randomised trial,¹ but has stabilised during the 2010s with about 6 million users in the USA and UK alone (figure 1, appendix p 4).

Current hormone therapy guidelines vary in what is said about ovarian cancer. The European drug regulatory guidelines² do not mention the disease, nor does the US Food and Drug Administration statement³ (based just on the Women's Health Initiative, which recorded few ovarian cancers). UK drug regulatory guidelines⁴ state that ovarian cancer might be increased by long-term use, but were dominated by findings from one large study;⁵ new UK guidelines are being developed. The most recent WHO review was completed before results from most large studies were published, so merely concluded that there was insufficient evidence about any ovarian cancer risk.⁶ Recently, some non-governmental reviews have argued that a few years of hormone therapy use starting before the age of 60 years

should cause no material harm.⁷ Most individual studies have, however, been too small to assess reliably any risks associated with use for only a few years (which is nowadays the usual pattern), so a systematic review of the worldwide epidemiological evidence is needed.

Reliable epidemiological assessment of any association of hormone therapy use with ovarian cancer requires large numbers and careful control of all potential sources of appreciable bias, and reviews just of the published evidence cannot provide this. For, although many studies of ovarian cancer collected some information about hormone therapy use, some were focused chiefly on other issues. Hence, published data about hormone therapy use are available for only about half the studies of ovarian cancer that have relevant data (appendix pp 5–9). Moreover, in some of the studies with retrospective designs, hormone therapy users might have been more willing than non-users to participate as controls, or there might have been differential recall of hormone therapy use between women already diagnosed with ovarian cancer and unaffected women. The Collaborative Group on

Lancet 2015; 385: 1835–42

Published Online
February 13, 2015
[http://dx.doi.org/10.1016/S0140-6736\(14\)61687-1](http://dx.doi.org/10.1016/S0140-6736(14)61687-1)
See [Editorial](#) page 1804
See [Comment](#) page 1806

*Analysis and writing committee listed at end of paper

Correspondence to:
Secretariat, Cancer Epidemiology Unit, NDPH, Richard Doll Building, Oxford OX3 7LF, UK
collaborations@ceu.ox.ac.uk

See [Online](#) for appendix

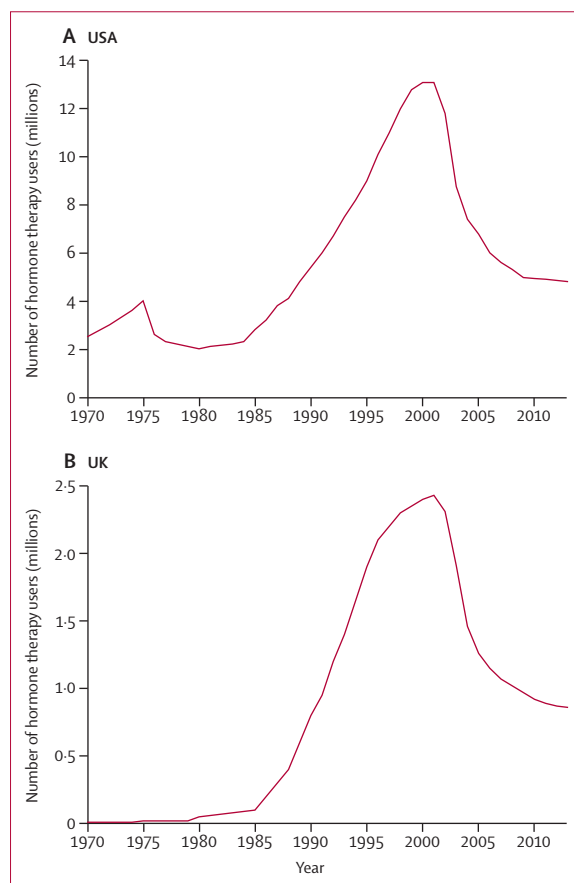


Figure 1: Trends in hormone therapy use in the USA and the UK since 1970
For source of data, see appendix p 4.

Epidemiological Studies of Ovarian Cancer was established in 1998 to bring together and analyse centrally individual participant data from all epidemiological studies of ovarian cancer, assessing the risks associated with hormonal and other factors. To evaluate with minimal bias the association of ovarian cancer with just a few years of hormone therapy use, or with past use, the principal analyses review detailed data from those prospective studies with information about both duration and recency of hormone therapy use. Sensitivity analyses review the evidence from all studies, prospective or retrospective.

Methods

Identification of studies and collection of data

Since 1998, epidemiological studies, published and unpublished, have been sought regularly by computer-aided literature searches, manual searches of review articles, written communications, and discussions at scientific meetings (see appendix p 5 for search strategy). Eligible studies are those with information on hormone therapy use, parity, oophorectomy and hysterectomy, and, if completed after 2006, at least 200 cases of ovarian cancer. By January, 2013, 58 such studies were identified

and principal investigators from each eligible study had been invited to collaborate. Datasets from 52 studies are included in these analyses and publications from three others are reviewed (appendix pp 5–10).

Cases are postmenopausal women with malignant, or borderline-malignant, epithelial or non-epithelial ovarian cancer; controls are postmenopausal women without ovarian cancer or previous oophorectomy.

In prospective studies, up to four randomly selected matched controls per case were selected. Individual data from 51 of the 52 studies were analysed centrally as case-control comparisons; because of data protection laws, individual data from the Danish Sex Hormone Register Study⁸ could not be exported, so its investigators provided detailed tabular results to combine with those of other studies (appendix p 6).

Information was sought for each woman on socio-demographic, reproductive, and other factors, including hormone therapy use before cancer diagnosis for cases and to an equivalent time for controls. Postmenopausal was defined as having reached natural menopause or age 55 years (because >90% of women have a natural menopause before that age⁹). Hysterectomy can mask natural menopause, so women younger than 55 years with a hysterectomy were excluded. Information sought about hormone therapy included ever-use, current use, age at first and last use, total duration of use, and constituents of each preparation. The hormone therapy preparation last used was classified as oestrogen-only or oestrogen-progestagen (or other/unknown formulation; appendix p 6).

Tumour histology was classified as fully malignant or borderline-malignant, and as epithelial or not. Epithelial tumours were further subdivided into the four most common types: serous, endometrioid, mucinous, or clear-cell (or mixed/other; appendix p 6). When appropriate, the International Classification of Diseases for Oncology¹⁰ was used.

Full details about information sources, search strategy, data collection, and definitions are provided in the appendix.

Statistical analyses and presentation of results

A protocol was circulated to collaborators and preliminary results were discussed at a meeting of investigators in July, 2011. Poisson logistic regression comparing particular groups of hormone therapy users with never-users yielded odds ratios, described here as relative risks (RRs). When more than two groups were compared (eg, current-or-recent users, long-term ex-users, and never-users), the variance of the log risk was estimated for each group (appendix p 6).¹¹ These group-specific variances were used to calculate group-specific CIs, facilitating valid comparisons between any two or more groups, whether or not one of them was designated as the baseline group.

The principal analyses include the prospective studies only, to avoid any possible biases associated with

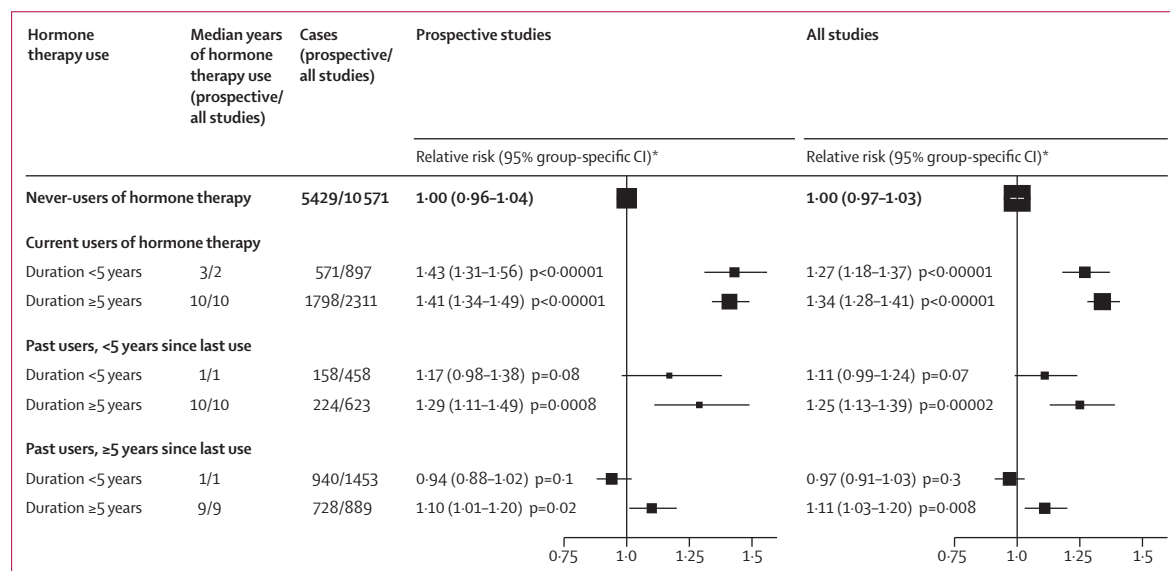


Figure 2: Relative risk of ovarian cancer by duration of use in current and past users of hormone therapy

*Risk relative to never-users of hormone therapy, stratified by age at diagnosis, study, and body-mass index, and adjusted for age at menopause, hysterectomy, oral contraceptive use, and parity. p values are two-sided and include the effects of the group-specific variance in never-users.

differential participation or recall in retrospective studies, but throughout the main report sensitivity analyses are given that include both the prospective and the retrospective studies. Results for the retrospective studies only are given in the appendix (pp 17, 19, 20), and heterogeneity tests were done to compare results from prospective and retrospective studies.

Because women can change their use of hormone therapy over time, follow-up in prospective studies was censored 4 years after hormone therapy use was last recorded (sensitivity analyses explored other cutoffs); duration and recency of use were estimated as if the last recorded use had continued (ie, duration of use in those who were current users when last asked increased by 1 year for each year of follow-up, as did time since last use of hormone therapy in ex-users). Hence, if in a prospective study of hormone therapy use the information last recorded before diagnosis is correct, then in analyses that combine current users with recent ex-users (ie, women who stopped <5 years before diagnosis), the current-or-recent users would include no misclassified women. Likewise, the never-users are contaminated with few hormone therapy users: only those who started in the interval of less than 4 years before diagnosis, which would be so uncommon (appendix p 6) as to dilute only slightly any real effects of hormone therapy use on risk.

To ensure that women in one study were compared directly only with similar women in that same study, all analyses were stratified by study, centre within study, age (5-year age groups up to 85–89 years), and body-mass index (<25, 25–29, or ≥30 kg/m²), and were adjusted for parity (0, 1–2, or ≥3), past use of oral contraceptives (never, <5 years of use, or ≥5 years of use), and age at menopause

(natural menopause before age 50 years, natural menopause after age 50 years, or previous hysterectomy). Unknowns for each variable were assigned to separate strata. Sensitivity analyses investigated additional adjustment for eight other potential confounding factors. Standard χ^2 tests for heterogeneity were used.

Results were weighted by the amount of statistical information in each stratum (inverse of the variance of log RR) and are presented as squares and lines, representing RRs and CIs (or, where appropriate, group-specific CIs). Study-specific results give 99% CIs (to allow for multiple testing), but most other results in the figures and all results in the text have 95% CIs. To illustrate the correspondence between relative and absolute risks in hormone therapy users, absolute risks were estimated from ovarian cancer incidence rates in England (appendix p 11).¹² Analyses were done with STATA 13.

Role of the funding sources

The study funders had no role in study design, data collection, analysis or interpretation, report preparation, or the decision to publish. The analysis and writing committee had full access to all the data and had final responsibility for the decision to submit for publication.

Results

Overall information was provided for 21488 postmenopausal women with ovarian cancer (cases) from 52 studies (17 prospective and 35 retrospective; appendix pp 7–10). The prospective studies contributed more than half of the cases (12110), with mean diagnosis year 2001 (SD 6), 55% (6601) of whom had used hormone therapy, with median duration 6 years

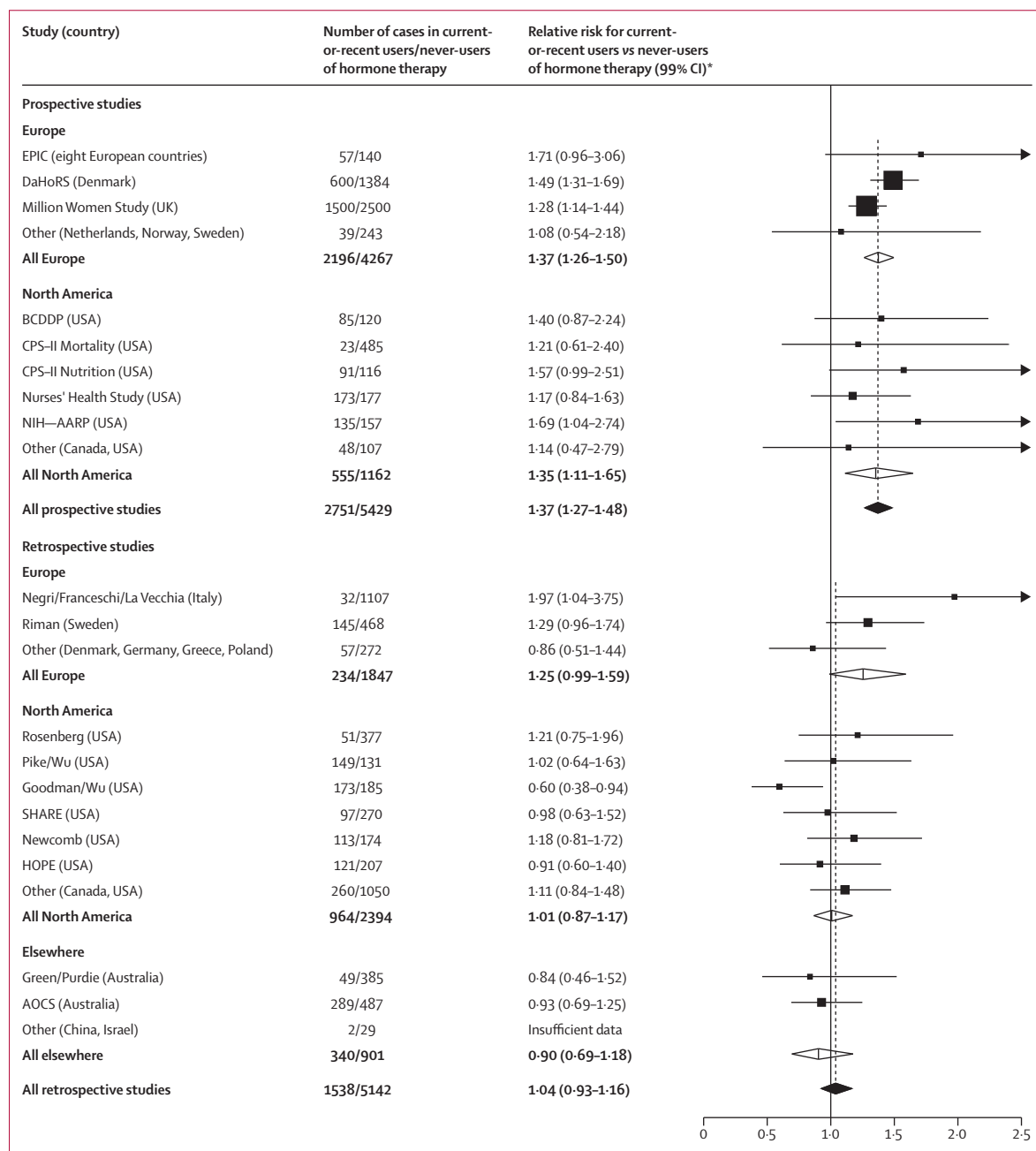


Figure 3: Study-specific results for the relative risk of ovarian cancer for current-or-recent users versus never-users of hormone therapy

For study-specific details and references, see appendix pp 7–10. Dotted lines represent totals for all prospective studies and, separately, for all retrospective studies. Study-specific results are arranged by study design and region; results are given for individual studies with the most statistical information (ie, with variance of log relative risk <0.03). Results for the remaining studies are grouped together here (and given separately for every study in appendix p 14). In comparisons of relative risks in prospective versus retrospective studies, overall heterogeneity $p < 0.0001$; for European studies, heterogeneity $p = 0.4$; and for North American studies, heterogeneity $p = 0.002$. In a comparison of relative risks in prospective studies, Europe versus North American heterogeneity $p = 0.9$; for retrospective studies, Europe versus North American heterogeneity $p = 0.04$. References provided in the appendix. *Risk relative to never-users of hormone therapy, stratified by age at diagnosis, study, and body-mass index, and adjusted for age at menopause, hysterectomy, oral contraceptive use, and parity.

(IQR 2–10). By contrast, in the retrospective studies only 29% (2702) of the women had used hormone therapy, with median duration 4 years (IQR 1–10), and the mean diagnosis year was 1992 (SD 8), well before peak hormone therapy use (figure 1).

Ovarian cancer risk was significantly greater in ever-users than in never-users of hormone therapy (RR 1.20, 95% CI 1.15–1.26, $p < 0.0001$ for prospective studies; 1.14, 1.10–1.19, $p < 0.0001$ for all studies combined; every study-specific result is provided in

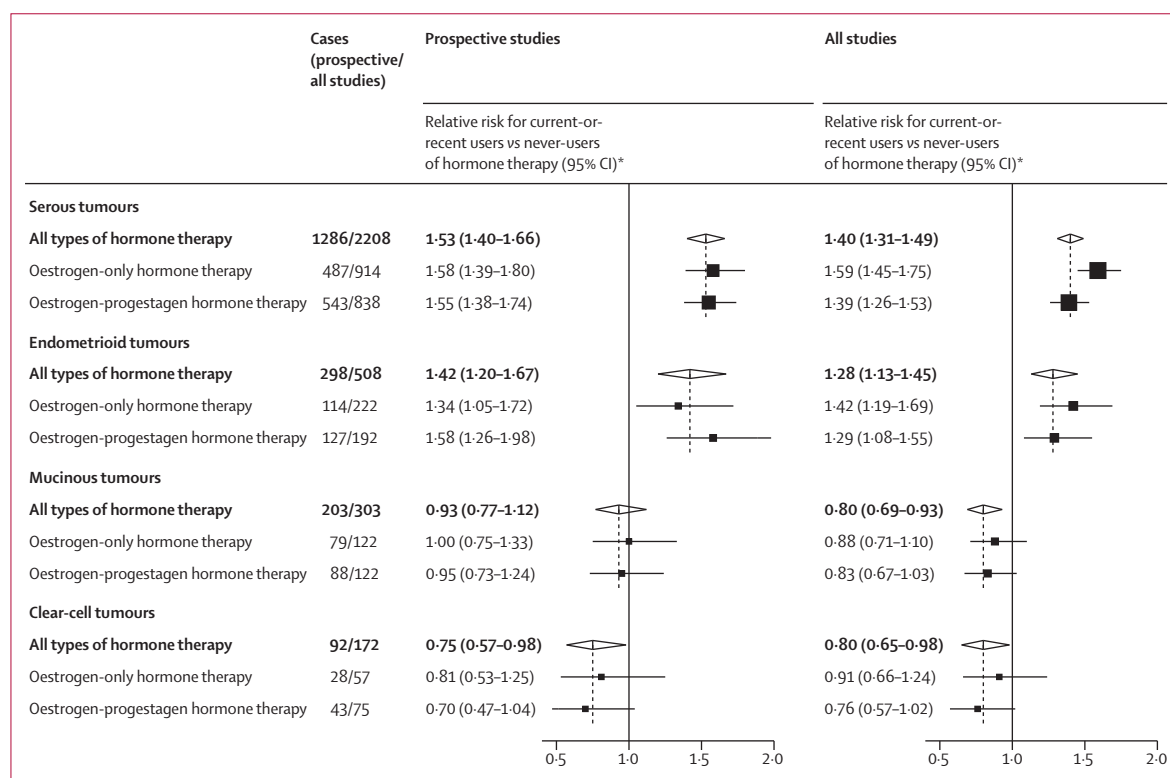


Figure 4: Relative risk of the four most common subtypes of ovarian cancer in current-or-recent users versus never-users of hormone therapy

Numbers do not add to totals, because some hormone therapy users were classified as using other or unknown types of hormone therapy and some epithelial tumours are classified as mixed types, other type, or type not specified. *Risks relative to never-users of hormone therapy, stratified by age at diagnosis, study, and body-mass index, and adjusted for age at menopause, hysterectomy, oral contraceptive use, and parity.

appendix p 12). Subsequent analyses were restricted to women with information both on duration of use and on time since last use of hormone therapy; this exclusion of studies without information on duration of use or time since last use slightly increased these RRs (appendix p 13).

Risk was strongly related to recency of use (figure 2). In prospective studies, risk was greatest in women who when last asked had been current users (RR 1.41, 95% CI 1.32–1.50; $p<0.0001$). Among them, risk was substantial even in those who, at diagnosis, had less than 5 years (median duration 3 years) of hormone therapy use (RR 1.43, 95% CI 1.31–1.56; $p<0.0001$).

Risk was, however, also significantly increased in women who had been recent ex-users and would at the time of diagnosis have still have been within 5 years of last use (RR 1.23, 95% CI 1.09–1.37; $p=0.0006$ in prospective studies). Risk decreased the longer ago hormone therapy had last been used, although women who had used hormone therapy for at least 5 years (median duration 9 years) and then stopped were still at significantly increased risk more than 5 years (median time since last use 10 years) later (RR 1.10, 95% CI 1.01–1.20; $p=0.02$). For prospective and retrospective studies combined, the risks were similar to those in prospective studies alone, except that the risks in

current users seemed to be somewhat smaller (figure 2, appendix p 17).

In prospective studies the risk for current-or-recent hormone therapy use (ie, use within the past 5 years) was 1.37 (95% CI 1.27–1.48; figure 3). This RR did not differ significantly between European and North American studies (1.37 vs 1.35; heterogeneity $p=0.9$).

However, the RR for current-or-recent hormone therapy use did differ significantly between prospective and retrospective studies (1.37 vs 1.04; heterogeneity $p<0.0001$; figure 3, appendix p 14). This difference was due to the lack of apparent effect in the aggregated North American retrospective studies (figure 1, appendix pp 12–14). In these retrospective studies, however, the design might have made unbiased recruitment of controls difficult, and the average year of diagnosis for the cases was 1990, well before hormone therapy use had become common (figure 1).

Sensitivity analyses left the main findings in prospective studies largely unchanged (appendix p 18). For example, adjustment for eight additional factors (year of birth, ethnic origin, education, age at menarche, height, alcohol consumption, smoking, and family history of ovarian or breast cancer) altered the RRs in current-or-recent users by 0.02 or less (the main findings had already been stratified by study, age, and body-mass index, and

	5 year incidence of ovarian cancer per 1000 never-users of hormone therapy	Absolute 5 year excess incidence per 1000 users with 5 years of hormone therapy use	Absolute 5 year excess incidence per 1000 users with 10 years of hormone therapy use
Age 50–54 years	1.2	0.52	0.52
Age 55–59 years	1.6	0.37	0.67
Age 60–64 years	2.1	0.10	0.61
Excess incidence	..	0.99 per 1000; 1 in 1000 users	1.80 per 1000; 1 in 600 users
Excess deaths	..	0.6 per 1000; 1 in 1700 users	1.2 per 1000; 1 in 800 users

Methods and sources of data are provided in appendix p 11.

Table: Estimated excess incidence of ovarian cancer in England associated with 5 years and with 10 years of hormone therapy use, starting at age 50 years

adjusted for parity, use of hormonal contraceptives, age at menopause, and hysterectomy).

Furthermore, the main findings were robust against variation in follow-up duration. Censoring either earlier or later than year 4 (appendix p 18) made little difference to the results. Two prospective studies (in the UK⁵ and Denmark⁸) contributed the most statistical information, but the RR in all other prospective studies was much the same when these two were excluded. One prospective study¹³ used fatal ovarian cancer as the outcome, but again its findings were typical.

In current-or-recent users, ovarian cancer risk was significantly increased with use of both oestrogen-only and oestrogen-progestagen preparations, with little heterogeneity between the risks: RR 1.37 (95% CI 1.26–1.50) and 1.37 (1.26–1.48), respectively, in the prospective studies (heterogeneity $p=0.9$); and 1.32 (1.23–1.41) and 1.25 (1.16–1.34) in all studies (heterogeneity $p=0.3$; appendix p 17). Few women had used both these classes of hormone therapy, and within the two classes there was insufficient information to assess whether risk varied by hormone therapy formulation or mode of delivery (appendix p 6).

All but three studies provided tumour histology (appendix p 6). Of tumours with known histology, 98% (14862 of 15090) were epithelial. The RR in current-or-recent users versus never-users did not seem to differ between epithelial (RR 1.28, 95% CI 1.22–1.34) and non-epithelial (1.35, 0.90–2.02) tumours, although the confidence interval for non-epithelial tumours was wide. There were only 228 non-epithelial tumours, too few for further analysis.

Almost all epithelial tumours of known, unmixed histology were adenocarcinomas of four main tumour types which were, in decreasing order of frequency, serous (7406 cases), endometrioid (1749), mucinous (1434), and clear cell (766). These four epithelial tumour types had qualitatively different relationships with hormone therapy use, both in prospective studies alone and in all studies combined (figure 4; heterogeneity $p<0.0001$).

In prospective studies, risks in current-or-recent users were definitely increased only for the two most common tumour types, serous (RR 1.53, 95% CI 1.40–1.66, $p<0.0001$) and endometrioid (1.42, 1.20–1.67; $p<0.0001$). In the aggregate of these two types, the risk more than 5 years since last use for past users who had used hormone therapy for more than 5 years (RR 1.25, 95% CI 1.07–1.46; $p=0.005$) was more definite than in the aggregate of all types (1.10, 1.01–1.20, $p=0.02$). Risk might have been somewhat decreased for the least common type, clear-cell tumours (RR 0.75, 95% CI 0.57–0.98; $p=0.04$ before any allowance for multiple hypothesis testing), but this protective effect is not statistically definite, and in the aggregate of both of the less common types the risk reduction was not significant (RR 0.86, 0.74–1.01; $p=0.07$). Within each tumour type there was little difference between the RRs for oestrogen-only and oestrogen-progestagen preparations (figure 4), or for borderline and fully malignant tumours (appendix p 19).

Age at initiation of hormone therapy had little effect; the RRs in current-or-recent users were similarly elevated with hormone therapy use beginning before age 50 years (RR 1.35, 95% CI 1.24–1.47) and at age 50–59 years (1.31, 1.22–1.40), with little information about older ages (1.15, 0.93–1.43; appendix p 20). Likewise, the available evidence suggested no major heterogeneity across subgroups defined by smoking, body size, parity, past use of oral contraceptives, hysterectomy, or other characteristics (appendix pp 15–16).

Application of the RRs in the prospective studies to age-specific ovarian cancer incidence and death rates in England suggested that 5 years of hormone therapy use, starting at around 50 years of age, would result in about one additional ovarian cancer per 1000 users and one additional ovarian cancer death per 1700 users (appendix p 11); 10 years of hormone therapy use from around 50 years of age would result in about one additional ovarian cancer per 600 users and one additional ovarian cancer death per 800 users (table).

Discussion

This collaboration brought together and analysed centrally individual data from 52 epidemiological studies, in which about half the postmenopausal women with ovarian cancer had used hormone therapy. Ovarian cancer risk was significantly increased in current users, even in those with less than 5 years hormone therapy use. In ex-users, risks decreased the longer ago hormone therapy use had ceased, but risks during the first few years after stopping remained appreciable. Furthermore, about a decade after ceasing long-duration hormone therapy use, there still seemed to be a small excess risk.

In current-or-recent users (all of whom had used hormone therapy within the past 5 years), the RRs did not differ significantly between users of oestrogen-only and of oestrogen-progestagen preparations, or between

women who had started hormone therapy before the age of 50 years or during their 50s. The RR did, however, vary substantially by tumour type, being increased only for the two most common histological types, serous and endometrioid tumours. In analyses restricted to these two types, the excess risk about a decade after ceasing long duration hormone therapy use became more definite.

An important strength of prospective studies is that recruitment takes place and information about hormone therapy use is recorded before women know whether they will develop ovarian cancer. The robustness of prospective data is demonstrated by the stability of the findings in various sensitivity analyses, and by the similarity of the findings in Europe and North America. Prospective studies provided more than half the statistical information, so results for all studies (prospective and retrospective combined) were broadly similar to those for the prospective studies alone.

When the retrospective studies were assessed in isolation, their aggregate findings differed from those of the prospective studies, however, perhaps because of biases in some retrospective studies. Many retrospective study results could have been somewhat biased by selective participation of hormone therapy users, and in all retrospective studies information about hormone therapy use was recorded after cancer diagnosis, so there might have been differential recall of hormone therapy use. Moreover, some retrospective datasets in this collaboration have yielded apparently discrepant findings on the association of ovarian cancer risk with smoking¹⁴ and with body-mass index.¹⁵

Almost all the worldwide evidence from eligible epidemiological studies was included in this meta-analysis. The three eligible studies that had published results but did not contribute data were all retrospective and North American, and had all reported increased risks of ovarian cancer associated with some aspect of hormone therapy use (appendix p 5). Had they been included, the findings in North American retrospective studies might not have been as different from those in other groups of studies.

As long as current users and recent ex-users are combined together as current-or-recent users, there are fewer potential sources of bias in prospective than in retrospective studies, and there is now for the first time sufficient evidence from prospective studies alone for statistically stable meta-analyses. Hence, it is now possible to base the main conclusions on prospective study results.

The overall relative risk for any type of ovarian cancer is the key public health outcome. There are, however, four main types of ovarian cancer, and in the prospective studies risk was definitely increased only for the two most common types, serous and endometrioid. Risk was possibly, although not definitely, decreased for the least common type, clear-cell tumours. Although

tumour histology could have been classified in slightly different ways in different studies, any misclassification would tend to blur differences by tumour type, yet we noted distinctly heterogeneous RRs. This heterogeneity argues strongly for causality, because it implies that the hormone-therapy-associated risks were not due just to confounding and that different ovarian cancer types have differing causes. The reasons for this heterogeneity are unclear, partly because the sites of origin of the four main tumour types are uncertain.¹⁶ The dependence of risk on ovarian tumour type is quite different for other exposures; oral contraceptives decrease serous, endometrioid, and clear-cell but not mucinous tumours,¹⁷ whereas smoking decreases endometrioid and clear-cell but increases mucinous tumours.¹⁴

The findings that ovarian cancer risk is greatest in current users of hormone therapy, falls after use ceases, and varies by tumour type, strongly suggest a causal relationship—ie, that among otherwise similar women, use of hormone therapy increases the probability of developing the two most common types of ovarian cancer, and hence ovarian cancer as a whole. There are still some 6 million users of hormone therapy in the USA and the UK, in addition to a comparable number in other high-income countries (figure 1, appendix p 4). At present, the WHO, European, and US guidelines about hormone therapy do not mention ovarian cancer, and the UK guidelines (which are due to be revised) state only that risk may be increased with long-term use. The definite risk of ovarian cancer that is observed even with less than 5 years of use starting at around age 50 years is directly relevant to current patterns of hormone therapy use, and hence directly relevant to medical advice, personal choices, and the current efforts to revise UK and worldwide guidelines.

Contributors

Analysis and writing committee: Cancer Epidemiology Unit, Oxford, UK: V Beral, K Gaitskell, C Hermon, K Moser, G Reeves; and Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Oxford, UK: R Peto. *Contributors:* The CEU secretariat (V Beral, K Gaitskell, C Hermon, K Moser, G Reeves) identified studies, received and checked data, conducted analyses, and had full access to all materials and results. V Beral, C Hermon, R Peto, and G Reeves drafted the report, and all writing committee members helped revise it before and after circulation to collaborators.

Collaborative Group on Epidemiological Studies of Ovarian Cancer

American Cancer Society, Atlanta, GA, USA: S M Gapstur, A V Patel; Australian National University, Canberra, ACT, Australia: E Banks; Aviano Cancer Center, Italy: L Dal Maso, R Talamini; Gertner Institute, Tel Hashomer, Israel: A Chetrit, G Hirsh-Yecheskel, F Lubin, S Sadetzki; Cancer Epidemiology Unit, Oxford, UK: V Beral, D Bull, B Cairns, B Crossley, K Gaitskell, A Goodill, J Green, C Hermon, T Key, K Moser, G Reeves; Cancer Council NSW, NSW, Australia: F Sitas; Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Oxford, UK: R Collins, R Peto; Catalan Institute of Oncology, Barcelona, Spain: C A Gonzalez; Centers for Disease Control and Prevention, Atlanta, GA, USA: N Lee, P Marchbanks, H W Ory, H B Peterson, P A Wingo; Chiang Mai University, Thailand: N Martin, S Silpisornkosol, C Theetranont; Chulalongkorn University, Bangkok, Thailand: B Boosiri, S Chutivongse, P Jimakorn, P Virutamasen, C Wongsrichanalai; Cedars-Sinai Medical Center, Los Angeles, CA,

USA: M T Goodman; Copenhagen University Hospital, Denmark: O Lidegaard, S K Kjaer, L S Mørch; Danish Cancer Society Research Center, Copenhagen, Denmark: S K Kjaer, A Tjønneland; Colorado School of Public Health, Denver, CO, USA: T Byers; Albert Einstein College of Medicine, Bronx, New York, NY, USA: T Rohan; Herlev University Hospital, Copenhagen, Denmark: B Mosgaard; Department of Public Health, Oxford, UK: M Vessey, D Yeates; Department of Social and Preventive Medicine, State University of New York, Buffalo, NY, USA: J L Freudenheim; Geisel School of Medicine at Dartmouth, Hanover, NH, USA: L J Titus; German Cancer Research Center (DKFZ), Heidelberg, Germany: J Chang-Claude, R Kaaks; University of Minnesota School of Public Health, Minneapolis, MN, USA: K E Anderson, D Lazovich, K Robien; Fred Hutchinson Cancer Research Center, University of Washington, Seattle, WA, USA: J Hampton, P A Newcomb, M A Rossing, D B Thomas, N S Weiss; Hillerød Hospital, University of Copenhagen, Denmark: E Lokkegaard; Imperial College London, UK: E Riboli; INSERM U1018 and Paris South University, Institut Gustave-Roussy, Villejuif, France: F Clavel-Chapelon; Harvard Medical School, Cambridge, MA, USA: D Cramer, S E Hankinson, R M Tamimi, S S Tworoger; International Agency for Research on Cancer, Lyon, France: S Franceschi; Istituto di Ricerche Farmacologiche Mario Negri, University of Milan, Italy: C La Vecchia, E Negri; Karolinska Institutet, Stockholm, Sweden: H O Adami, C Magnusson, T Riman, E Weiderpass, A Wolk; Maastricht University, Netherlands: L J Schouten, P A van den Brandt; Mahidol University, Bangkok, Thailand: N Chantarakul, S Koetsawang, D Rachawat; Cancer Research and Prevention Institute, Florence, Italy: D Palli; National Cancer Institute, Bethesda, MD, USA: A Black, L A Brinton, D M Freedman, P Hartge, A W Hsing, J V Lacey Jr, J Lissowska, R N Hoover, C Schairer; Cancer Epidemiology Research Group, National Health Laboratory Services, Medical Research Council, Johannesburg, South Africa: C Babb, M Urban; Norwegian Institute of Public Health, Oslo, Norway: S Graff-Iversen, R Selmer; QIMR Berghofer Medical Research Institute and University of Queensland, Brisbane, QLD, Australia: C J Bain, A C Green, D M Purdie, V Siskind, P M Webb; Roswell Park Cancer Institute, Buffalo, NY, USA: K Moysich, S E McCann; Royal College of General Practitioners' Oral Contraception Study, London, UK: P Hannaford, C Kay; School of Public Health, Curtin University, Perth, WA, Australia: C W Binns, A H Lee, M Zhang; School of Public Health, University of Texas, Houston, TX, USA: R B Ness; School of Public Health and Health Sciences, University of Massachusetts, Boston, MA, USA: P Nasca; Slone Epidemiology Center, Boston University, MA, USA: P F Coogan, J R Palmer, L Rosenberg; Stanford University, CA, USA: A Whittemore; University of Athens Medical School, Athens, Greece: K Katsouyanni, A Trichopoulou, D Trichopoulos, A Tzonou; University of Chile, Santiago, Chile: A Dabancens, L Martinez, R Molina, O Salas; University of Hawaii, Honolulu, HA, USA: G Lurie, M E Carney, L R Wilkens; University Hospital, Lund, Sweden: L Hartman, J Manjer, H Olsson; University Hospital of North Norway, Tromsø, Norway: M Kumle; University of Pennsylvania, Philadelphia, PA, USA: J A Grisso, M Morgan, J E Wheeler; University of Pittsburgh, PA, USA: R P Edwards, J L Kelley, F Modugno; University Medical Center, Utrecht, Netherlands: N C Onland-Moret, P H M Peeters; University of Southern California, Los Angeles, CA, USA: J Casagrande, M C Pike, A H Wu; University of New South Wales, NSW, Australia: K Canfell; University of Toronto, ON, Canada: A B Miller; Arctic University of Tromsø, Norway: I T Gram, E Lund; George Washington University, Washington, DC, USA: L McGowan; Vanderbilt University, Nashville, TN, USA: X O Shu, W Zheng; World Health Organization, Geneva, Switzerland: T M M Farley, S Holck, O Meirik; and Yale School of Public Health, CT, USA: H A Risch.

Declaration of interests

All analysis and writing committee members declare no competing interests.

Acknowledgments

The chief acknowledgment is to the participants and to the investigators in these studies. Data checking and analysis in the Oxford Cancer Epidemiology Unit was supported by Cancer Research UK and the UK Medical Research Council. Adrian Goodill drew the graphs.

References

- 1 Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002; **288**: 321–33.
- 2 European Medicines Agency. Guidelines on clinical investigation of medicinal products for hormone replacement therapy of oestrogen deficiency symptoms in postmenopausal women www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003348.pdf (accessed Jan 1, 2014).
- 3 US Food and Drug Administration. Estrogen and estrogen with progestin therapies for postmenopausal women www.fda.gov/Drugs/DrugSafety/InformationbyDrugClass/ucm135318.htm (accessed Jan 1, 2014).
- 4 Medicines and Healthcare products Regulatory Agency (MHRA). Hormone-replacement therapy: safety update. UK Public Assessment Report www.mhra.gov.uk/home/groups/pl-p/documents/websiteresources/con2032228.pdf (accessed Jan 1, 2014).
- 5 Beral V, Reeves G, Green J, Bull D. Ovarian cancer and hormone replacement therapy in the Million Women Study. *Lancet* 2007; **369**: 1703–10.
- 6 International Agency for Research on Cancer. Monograph Volume 91: combined estrogen-progestogen contraceptives and combined estrogen-progestogen menopausal therapy. Lyon, France: International Agency for Research on Cancer, 2007.
- 7 de Villiers TJ, Pines A, Panay N, et al. Updated 2013 International Menopause Society recommendations on menopausal hormone therapy and preventive strategies for midlife health. *Climacteric* 2013; **16**: 316–37.
- 8 Mørch LS, Løkkegaard E, Andreasen AH, et al. Hormone therapy and ovarian cancer. *JAMA* 2009; **302**: 289–305.
- 9 Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncology* 2012; **13**: 1141–51.
- 10 Fritz AG, Percy C, Jack A, et al, eds. International classification of diseases for oncology (ICD-O). Geneva: World Health Organization, 2000.
- 11 Plummer M. Improved estimates of floating absolute risk. *Stat Med* 2004; **23**: 93–104.
- 12 Office for National Statistics. National Cancer Statistics for England. <http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Cancer> (accessed Jan 1, 2014).
- 13 Rodriguez C, Patel AV, Calle EE, Jacob EJ, Thun MJ. Estrogen replacement therapy and ovarian cancer mortality in a large prospective study of US women. *JAMA* 2001; **285**: 1460–65.
- 14 Collaborative Group on Epidemiological Studies of Ovarian Cancer. Ovarian cancer and smoking: individual participant meta-analysis including 28 114 women with ovarian cancer from 51 epidemiological studies. *Lancet Oncol* 2012; **13**: 946–56.
- 15 Collaborative Group on Epidemiological Studies of Ovarian Cancer. Ovarian cancer and body size: individual participant meta-analysis including 25 157 women with ovarian cancer from 47 epidemiological studies. *PloS Med* 2012; **9**: e1001200. DOI:10.1371/journal.pmed.1001200.
- 16 Kurman RJ, Shih IeM. Molecular pathogenesis and extra-ovarian origin of epithelial ovarian cancer—shifting the paradigm. *Hum Pathol* 2011; **42**: 918–31.
- 17 Collaborative Group on Epidemiological Studies of Ovarian Cancer. Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies including 23 257 women with ovarian cancer and 87 303 controls. *Lancet* 2008; **371**: 303–14.

APPENDIX D. PAPERS PUBLISHED DURING DPHIL AS PART OF COLLABORATIONS WITH OTHERS

D.2 PAST CERVICAL INTRAEPITHELIAL NEOPLASIA GRADE 3, OBESITY, AND EARLIER MENOPAUSE ARE ASSOCIATED WITH AN INCREASED RISK OF VULVAL CANCER IN POSTMENOPAUSAL WOMEN (BRITISH JOURNAL OF CANCER, 2016)

This paper resulted from a collaboration with another DPhil student at CEU (Dr Kate Coffey), and I am listed as the second author. I contributed to the histological classification and to discussions of the results, and I read and commented on drafts of the paper.

Keywords: vulval cancer; obesity; CIN 3; Million Women Study; postmenopausal

Past cervical intraepithelial neoplasia grade 3, obesity, and earlier menopause are associated with an increased risk of vulval cancer in postmenopausal women

Kate Coffey¹, Kezia Gaitskell¹, Valerie Beral¹, Karen Canfell^{2,3}, Jane Green¹, Gillian Reeves¹ and Isobel Barnes^{*,1} on behalf of the Million Women Study Collaborators⁴

¹Cancer Epidemiology Unit, University of Oxford, Richard Doll Building, Old Road Campus, Roosevelt Drive, Oxford, OX3 7LF, UK;

²Cancer Research Division, Cancer Council NSW, Woolloomooloo, New South Wales 2011, Australia and ³School of Public Health, Sydney Medical School, University of Sydney, Sydney, New South Wales, Australia

Background: Vulval cancer predominantly affects postmenopausal women. A smaller proportion of vulval cancers, particularly at older ages, are now thought to be associated with human papillomavirus infection than previously reported, but other risk factors have not been well examined in prospective cohort studies.

Methods: A total of 1.3 million women aged 49–65 years were followed for incident vulval cancer (ICD-10 C51). Adjusted Cox regression models were used to examine the relationship between reproductive and lifestyle factors and risk of vulval cancer.

Results: There were 898 vulval cancers registered in the cohort over an average of 14 years of follow-up; 70% were squamous cell carcinomas. Past registration of cervical carcinoma *in situ* (RR 2.68; 95% CI 1.71–4.18; $P < 0.001$), obesity (RR 1.71; 95% CI 1.44–2.04; $P < 0.0001$), and menopause before the age of 50 years (RR 1.52; 95% CI 1.22–1.89; $P < 0.001$) were associated with a significantly increased risk of subsequent vulval cancer.

Conclusion: Past cervical pre-cancer, obesity, and earlier age at menopause are associated with an increased risk of vulval cancer at older ages.

Although vulval cancer accounts for <1% of all female cancer, it is the most common anogenital cancer in women over 70 years (Figure 1). In 2012, there were 1052 vulval cancers registered in England, of which 86% (909 cases) were in women aged over 50 years. Its annual incidence increases markedly with age, rising from 2.5 cases per 100 000 women at ages 40–44 years to 17.0 per 100 000 at ages 80–84 years (Office for National Statistics, 2014).

It has been shown that the risk of vulval cancer is associated with cervical intraepithelial neoplasia (CIN) and that vulval cancer has aetiopathological similarities with high-risk human

papillomavirus (HrHPV)-related cancers such as cervical, anal, and vaginal cancers (Kalliala *et al*, 2005; Crawford *et al*, 2011). Until recently, it was thought that ~40% of vulval cancer cases were attributable to HrHPVs (Chaturvedi, 2010). However, it now seems that a decreasing proportion of vulval cancers are associated with HrHPV as women age; it has been reported that only 15–17% of cases in women aged ≥67 years are associated with HrHPV infection (de Sanjosé *et al*, 2013). Vulval cancers in older women are thought to be more likely to arise in a background of vulval dermatosis (van de Nieuwenhof *et al*, 2009), and less likely to be

*Correspondence: Dr I Barnes; E-mail: isobel.barnes@ceu.ox.ac.uk

⁴Members of the Million Women Study Collaborators are listed before References.

Received 10 November 2015; revised 8 April 2016; accepted 30 April 2016

© 2016 Cancer Research UK. All rights reserved 0007–0920/16

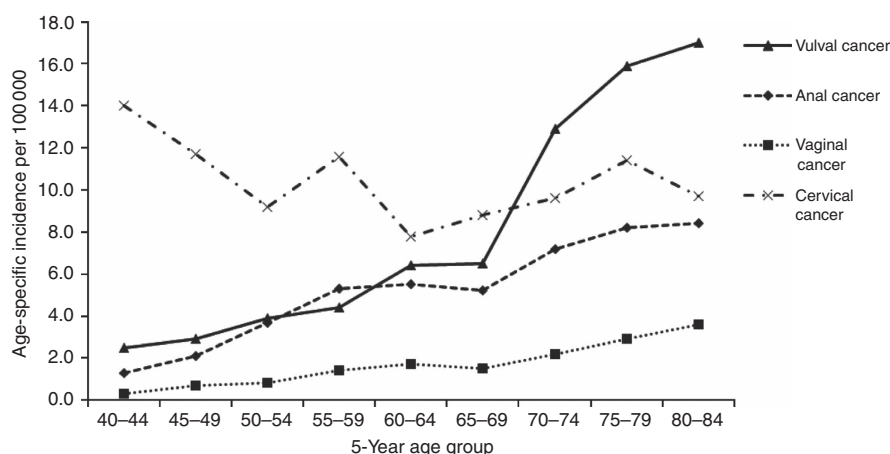


Figure 1. Age-specific rates of vulval cancer (ICD-10 C51) compared with cervical (C53), anal (C21), and vaginal (C52) cancer, England 2012. Rates are per 100 000 women by 5-year age groups, starting at age 40 years (Office for National Statistics, 2014).

HrHPV driven. In addition, cohort period and age effects relating to HrHPV exposure are likely to play a role (Barlow *et al*, 2015).

Most previous epidemiological studies of risk factors have been small and no consistent findings have emerged (Japaze *et al*, 1977; Newcomb *et al*, 1984; Brinton *et al*, 1990; Parazzini *et al*, 1993; Madsen *et al*, 2008). Some uncontrolled case series suggested a relationship between obesity and vulval cancer risk (Green *et al*, 1958; Franklin and Rutledge, 1972); however, a 2013 review concluded that although there may be a relationship between adiposity and vulval cancer, studies had been too small to answer the question definitively (Webb, 2013).

We investigated risk factors for vulval cancer in a large cohort of UK women aged ≥ 50 years, examining associations for vulval cancer overall, and by histological subtype, with obesity, diabetes, and various lifestyle and reproductive risk factors including smoking, parity, past cervical dysplasia, age at menopause, and use of oral contraceptives and menopausal hormone therapy (HT).

MATERIALS AND METHODS

The Million Women Study is a large prospective cohort study that recruited 1.3 million women invited for routine breast screening via the National Health Service Breast Screening Programme (NHSBSP) between 1996 and 2001 (i.e. women who were born between 1931 and 1952). Full details of the study design and methods are described elsewhere (Beral *et al*, 2003). Study questionnaires and further details of the data and access policies can be viewed online at www.millionwomenstudy.org.

Study participants completed a baseline questionnaire providing self-reported information on lifestyle, health and reproductive exposures. Participants were resurveyed twice, ~ 4 and 8 years after recruitment. The study has Multi-Centre Research Ethics Committee approval (MREC 97/01).

The study teams are provided with regularly updated information on incident cancers and deaths via the UK National Health Service Central Registers (NHSCR). Cases of incident vulval cancer were identified as 'malignant neoplasm of vulva', International Classification of Diseases, version 10 (ICD-10) (World Health Organization 2011) C51. We used self-reported data from the recruitment questionnaire to define most exposures, including menopausal HT use. Excellent agreement has been found between HT use as reported by Million Women Study participants and general practice records in a past validation study (Banks *et al*, 2001).

Marital status and history of cervical screening were based on self-reported data from the first resurvey, when information on these exposures was first collected. The Townsend Deprivation Index (Townsend, 1987) was calculated using post-code and 1991 census data. Body mass index was calculated from self-reported weight and height measurements at recruitment that have been shown to correlate closely with measured variables in Million Women Study participants (Bobrow *et al*, 2012). Women with a registration of CIN grade 3 (CIN 3) before recruitment were identified using the information provided by NHSCR.

Vulval cancers were divided into four main histological subtypes: squamous cell carcinomas (SCCs), basal cell carcinomas (BCCs), glandular tumours, and melanocytic tumours. Histological subtypes were identified using ICDO-3 (International Classification of Diseases for Oncology version 3) morphology codes (World Health Organization, 2000; Kurman *et al*, 2014).

Squamous tumours included the following ICDO-3 codes: M8051/3, M8070/3, M8070/5, M8071/3, M8072/3, M8074/3, M8075/3; and M8076/3. Basal cell carcinomas included: M8090/3, M8091/3, M8094/3, and M8097/3. Glandular tumours (including Paget's disease, Bartholin gland tumours, and adenocarcinomas) were: M8140/3, M8200/3, M8390/3, M8400/3, M8480/3, M8542/3, and M8560/3. Finally, melanocytic tumours included: M8720/3, M8721/3, M8743/3, M8744/3, and M8746/3. Other rarer tumour types were excluded from the analysis, and these included unclassified malignancies and sarcomas (see the Results section for further enumeration of 'other' tumour histologies).

Statistical analyses. Participants were excluded from the analyses if they had any invasive cancer registered before the start of follow-up, other than non-melanoma skin cancer (ICD-10 C44). For analyses of all exposures apart from marital status and history of cervical screening, follow-up started on the date of recruitment or at the age of 50 years for the small number of women (3%) who were < 50 years old at recruitment. For marital status and history of cervical screening, follow-up started on the date that participants completed the first resurvey. Follow-up continued until the earliest of: the date of registration of any cancer; death; emigration or other loss to follow-up; or 31 December 2013, the end of follow-up.

We used Cox proportional hazard models with attained age as the underlying time variable to estimate hazard ratios (referred to as relative risks (RRs)) of developing vulval cancer in relation to the following exposures: smoking (never, past, current); alcohol (0–2, 3+ units per week); BMI (normal: $< 25 \text{ kg m}^{-2}$, overweight: $25\text{--}29 \text{ kg m}^{-2}$, obese: $30+ \text{ kg m}^{-2}$); diabetes (no, yes); oral contraceptive use (never, ever); parity (parous, nulliparous); prior hysterectomy (no, with oophorectomy, without oophorectomy);

prior CIN 3 registration (no, yes); deprivation (tertiles); marital status (married/living with a partner, or not); past cervical screening (no, yes); age at menopause (<50, 50+ years); and HT use (never, past, current).

All analyses were, where possible, mutually adjusted for exposures. In addition, all analyses were stratified by region of residence consisting of 10 geographical areas corresponding to the cancer registry regions: Oxford, East Anglia, South West, Thames, West Midlands, Northern and Yorkshire, Trent, North West (Mersey), North West (Manchester/Lancashire), and Scotland. The assumption of proportional hazards was examined using plots of the Schoenfeld residuals and was found to be reasonable.

In order to avoid distortion by HT use, analyses of age at menopause were restricted to women who reported at baseline that they had never used HT and that they either had a natural menopause or had a bilateral oophorectomy. Analyses of HT use were restricted to women who reported being post-menopausal at baseline.

A subgroup analysis was performed using a competing hazards model to examine the associations between BMI (<30, 30+ kg m⁻²), smoking (never, ever), age at menopause (<50, 50+ years), menopausal HT use (never, past, current), and vulval cancer by histological subtype. Results were tested for heterogeneity of effect of these exposures by tumour histological type. We were unable to look at the effect of CIN 3 by histology because of the small number of cases with this exposure within subtype groups.

We performed two sensitivity analyses. In the first, we excluded the first years of follow-up from the main analysis to assess the possibility of reverse association. In the second, we assessed the possibility of bias due to changes in exposures over the follow-up period. To do this we censored women who did not respond to a resurvey 4 years after their previous survey. We then compared an analysis in which exposures were fixed at recruitment to one in which smoking status, alcohol use, BMI, diabetes, and hysterectomy/oophorectomy were updated using self-reported data from the first and second resurveys, and CIN 3 registration was updated using data provided by the NHSCR.

The few women with missing data for adjustment variables (<5% for each variable) were included in a separate category for each variable of interest. Relative risks are reported with 95% confidence intervals (CIs). Analyses were performed in Stata 14 (Statacorp, 2015).

RESULTS

A total of 1.3 million women without prior cancer were followed for 18 million person-years (on average 13.8 years per woman), and there were 898 vulval cancers registered in the cohort over that period. Just over half of the cohort were overweight or obese: 34% of women had a BMI of 25 to 29.9 kg m⁻², putting them in the overweight range, and 17% were obese with a BMI of ≥30 kg m⁻². Most women were parous, with 89% of participants reporting ever having given birth. Nearly 60% had a history of oral contraceptive use, 46% had ever smoked, with 19% reporting being current smokers. Around half of all participants had ever used menopausal hormone therapy (Table 1).

The exposure with the strongest association with risk of vulval cancer later in life was a registration of CIN 3 before recruitment. Women with a CIN 3 registration before recruitment had more than a doubling in risk of vulval cancer compared with women who had not had CIN 3 (RR 2.68; 95% CI 1.71–4.18; *P*<0.0001).

Being overweight or obese was associated with an increased risk of vulval cancer compared with women with a normal BMI and risk increased with adiposity. Overweight women (BMI of

Table 1. Participant characteristics at baseline, and details of follow-up in the cohort

Characteristic at baseline	
Number of women (%)	1 300 042 (100.0)
Mean age at recruitment (s.d.)	56.6 (4.8)
Parous, % (n)	89.0 (1 157 188)
Oral contraceptive user (ever), % (n)	58.6 (761 700)
Current smoker, % (n)	19.2 (249 014)
Body mass index ≥30 kg m ⁻² , % (n)	17.0 (220 976)
Most deprived tertile, % (n)	33.0 (428 383)
Ever hormone therapy user, % (n)	49.6 (644 804)
Hysterectomy, % (n)	24.4 (316 927)
Registration of CIN 3 before recruitment, % (n)	1.0 (12 533)
Follow-up	
Women-years (1000s)	17 956
Mean length of follow-up per woman (s.d.)	13.8 (3.4)
Number of women with incident vulval cancer	898

Abbreviation: CIN 3 = cervical intraepithelial neoplasia grade 3.

25–29 kg m⁻²) had a 19% increase in risk of incident vulval cancer, with a RR of 1.19 (95% CI 1.02–1.39; *P*<0.0001), and this rose further in women classed as obese (BMI of ≥30 kg m⁻²), who had a RR of 1.71 (95% CI 1.44–2.04; *P*<0.0001; Figure 2). After adjustment for BMI and other factors, there was no significant increase in risk associated with a history of diabetes.

There was no association seen between smoking and overall vulval cancer risk; women who reported ever having smoked had a RR of 0.98 (95% CI 0.82–1.17) compared with never smokers. Similarly, no significant increase in risk was seen with alcohol use, age at menarche, oral contraceptive pill use, parity, or tubal ligation.

We examined risk associated with age at menopause in women who had never used menopausal HT and who reported a natural menopause or bilateral oophorectomy. Women with an earlier age at menopause had an increased risk of vulval cancer later in life. Menopause before age 50 years was associated with a 50% increase in risk of vulval cancer compared with menopause that occurred after 50 years (RR 1.52; 95% CI 1.22–1.89; *P*<0.001). There did not appear to be any further increase in risk associated with age at menopause before age 45 years (RR 1.39; 95% CI 1.00–1.94) compared with menopause at 45–49 years (RR 1.58; 95% CI 1.24–2.01). Use of menopausal hormone therapy was not associated with risk of subsequent vulval cancer overall (Table 2).

In analyses starting at the first resurvey at ~4 years after recruitment, when women were asked for the first time about marital status and attendance for cervical screening, there was no significant difference in risk of vulval cancer for women not currently married or living with a partner (RR 1.19; 95% CI 0.95–1.50) compared with those who were; or for those who reported having ever vs never attended for cervical screening (RR 0.96; 95% CI 0.57–1.62).

Two-thirds of vulval cancers (*n*=609, 68%) were specified as squamous cell carcinomas. Glandular tumours were less common, with 96 registered (11%), including Paget's disease and various types of adenocarcinoma. Basal cell tumours were the next most common, with 92 cases (10%). Melanocytic (*n*=55, 6%) and other (unspecified and sarcomatous) tumours made up the remaining number (*n*=46, 5%) (Table 3).

When we examined associations by histological subtype (Table 4), obese women (BMI of 30+ kg m⁻²) had a substantially and significantly elevated risk of squamous cell carcinoma (RR 1.84; 95% CI 1.53–2.21), but little or no associated increased risk for basal cell, glandular, or melanocytic tumours (test for heterogeneity *P*=0.01). We also observed a reduction in melanocytic tumours in women who had ever smoked (RR 0.49; 95% CI 0.27–0.91) compared with never smokers.

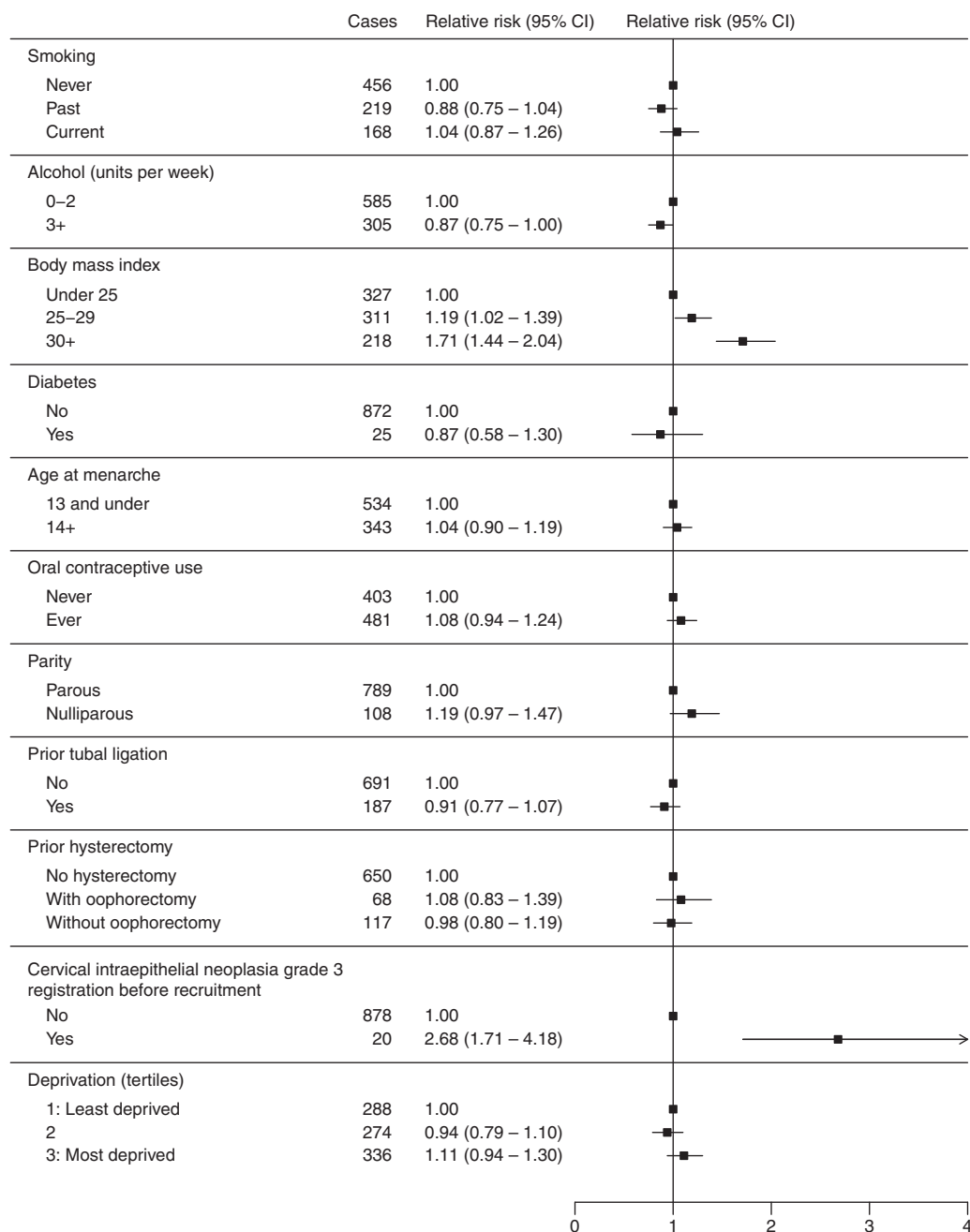


Figure 2. Association between vulval cancer and various lifestyle and other factors. Relative risks adjusted, where possible, for smoking, alcohol use, body mass index, diabetes, age at menarche, oral contraceptive use, parity, prior tubal ligation, prior hysterectomy, cervical intraepithelial neoplasia grade 3 before recruitment, and deprivation.

Table 2. Associations between risk of vulval cancer, age at menopause, and use of menopausal hormone therapy in postmenopausal women

Exposure	Cases/population at risk	Relative risk ^a	95% Confidence interval	P-value
Age at menopause in women with natural menopause or oophorectomy who have never used hormone therapy				
50+ Years	158/237 144	1.00		<0.001
< 50 Years	167/175 489	1.52	(1.22–1.89)	
Use of menopausal hormone therapy at baseline, all postmenopausal women				
Never	401/493 167	1.00		0.54
Past	142/192 727	0.94	(0.77–1.14)	
Current	252/424 544	0.86	(0.73–1.02)	

^aRelative risks adjusted, where possible, for smoking, alcohol use, body mass index, age at menarche, oral contraceptive use, parity, tubal ligation, cervical intraepithelial neoplasia grade 3, diabetes, hysterectomy/oophorectomy, and deprivation.

Table 3. Incident vulval cancer registrations in the cohort by histological subtype and International Classification of Diseases for Oncology, 3rd Edition (ICDO-3) morphology coding

ICDO-3 morphology code	Histological subtype	Number (% of all vulval cancers)
Squamous tumours		
M8070/3	Squamous cell carcinoma, not otherwise specified	447 (50.0)
M8071/3	Squamous cell carcinoma, keratinising, not otherwise specified	131 (14.6)
M8070/5, M8076/3	Squamous cell carcinoma, microinvasive	16 (1.8)
M8051/3	Verrucous carcinoma, not otherwise specified	8 (0.9)
M8072/3, 8074/3, 8075/3	Squamous cell carcinoma: large cell, non-keratinising/spindle cell/adenoid	7 (0.8)
	Total squamous cell tumours	609 (67.8)
Glandular tumours		
M8542/3	Paget's disease	73 (8.1)
M8480/3, M8400/3, M8140/3	Adenocarcinoma: mucinous/sweat gland/not otherwise specified	13 (1.4)
M8200/3, M8390/3, M8560/3	Adenoid cystic/skin appendage/adenosquamous carcinoma	10 (1.1)
	Total glandular tumours	96 (10.7)
Basal cell tumours		
M8090/3	Basal cell carcinoma, not otherwise specified	80 (8.9)
M8091/3	Multifocal superficial basal cell carcinoma	6 (0.7)
M8094/3, M8097/3	Basosquamous carcinoma, basal cell carcinoma (nodular)	6 (0.7)
	Total basal cell carcinomas	92 (10.2)
Melanocytic tumours		
M8720/3	Malignant melanoma, not otherwise specified	21 (2.3)
M8721/3	Nodular melanoma	16 (1.8)
M8743/3	Superficial spreading melanoma	14 (1.6)
M8744/3	Acrall lentiginous melanoma, malignant	4 (0.4)
M8746/3	Mucosal lentiginous melanoma	4 (0.4)
	Total melanocytic tumours	55 (6.1)
Other tumours: unspecified and sarcomatous		
M8000/3, 8010/3	Neoplasm/epithelial tumour, malignant	37 (4.0)
M8032/3, M8033/3, M8800/3, M8832/3, M8851/3, M8890/3, M8891/3	Spindle cell carcinoma, not otherwise specified/pseudosarcomatous carcinoma/sarcoma, not otherwise specified/dermatofibrosarcoma, not otherwise specified/liposarcoma, well-differentiated /leiomyosarcoma/epithelioid leiomyosarcoma	9 (0.9)
	Total 'other' tumours	46 (5.1)
	Total incident vulval cancers	898 (100.0)

Table 4. Associations between body mass index, smoking, age at menopause, menopausal hormone therapy use, and risk of vulval cancer by tumour histology subtype

Exposure	Squamous cell carcinoma			Glandular tumours			Basal cell carcinoma			Melanocytic tumours			P-value ^b
	Cases/ population at risk	RR ^a	95% CI	Cases/ population at risk	RR ^a	95% CI	Cases/ population at risk	RR ^a	95% CI	Cases/ population at risk	RR ^a	95% CI	
Body mass index													
<30	406/1 010 993	1.00		78/1 010 993	1.00		76/1 010 993	1.00		43/1 011 093	1.00		0.01
30+	168/220 976	1.84	(1.53–2.21)	15/220 976	0.87	(0.49–1.51)	16/220 976	1.03	(0.60–1.79)	11/220 976	1.20	(0.61–2.36)	
Smoking													
Never	294/627 145	1.00		47/627 145	1.00		52/627 145	1.00		36/627 145	1.00		0.06
Ever	276/596 249	1.04	(0.88–1.23)	46/596 249	1.13	(0.74–1.71)	34/596 249	0.75	(0.48–1.17)	15/596 249	0.49	(0.27–0.91)	
Age at menopause													
50+ Years	112/237 144	1.00		13/237 144	1.00		15/237 144	1.00		9/237 144	1.00		0.94
<50 Years	122/175 489	1.59	(1.22–2.06)	15/175 489	1.62	(0.76–3.45)	15/175 489	1.43	(0.69–2.97)	7/175 489	1.17	(0.42–3.23)	
Menopausal hormone therapy use													
Never	288/493 167	1.00		35/493 167	1.00		34/493 167	1.00		20/493 167	1.00		0.28
Past	87/192 727	0.84	(0.66–1.07)	20/192 727	1.51	(0.87–2.65)	20/192 727	1.69	(0.96–2.97)	6/192 727	0.79	(0.32–2.00)	
Current	159/424 544	0.82	(0.66–1.00)	33/424 544	1.26	(0.76–2.09)	25/424,54	1.15	(0.66–2.00)	24/424 544	1.53	(0.81–2.89)	

Abbreviations: CI = confidence interval; RR = relative risk.
^aRelative risks (RRs) adjusted where appropriate for smoking, alcohol use, body mass index, age at menarche, parity, tubal ligation, hysterectomy/oophorectomy, and deprivation.
^bP for heterogeneity.

The sensitivity analyses showed that our results are not likely to be due to reverse association (Supplementary Figure 1). They also showed that any bias caused by changes in the exposures over time is likely to be small (Supplementary Figures 2 and 3).

DISCUSSION

The strongest increase in risk of vulval cancer in our study was seen in women with a registration of CIN 3 before recruitment,

who had more than a doubling in risk compared with women who had not had CIN 3. Adiposity is also significantly associated with an increased risk of vulval cancer in women >50 years of age, principally in obese women, who have a 70% increase in vulval cancer incidence compared with those with a normal BMI; furthermore, the effect of BMI appears to be strongest for tumours with squamous cell morphology. Women who reported undergoing a natural or surgical menopause (oophorectomy) before age 50 years have a 52% increase in risk compared with those reporting a later menopause at ≥ 50 years.

Risk of vulval cancer was not significantly associated with smoking; alcohol use; marital status; parity; hysterectomy; ever having attended for cervical cancer screening; diabetes; prior tubal ligation; age at menarche; or use of exogenous hormones in the form of the oral contraceptive pill or menopausal HT.

The HrHPV-related anogenital cancers are known to be related to previous genital tract dysplasia, typically screen-detected disease of the cervix (Crawford *et al*, 2011). A Finnish retrospective cohort study (Kalliala *et al*, 2005) has previously looked at risk of vulval cancer following a diagnosis of any grade of CIN. They found a standardised incidence ratio of 4.1 (95% CI 1.5–8.9) for vulval cancer after treatment of CIN 1, 2, or 3. This was based on 6 cases of vulval cancer among 7564 women who had been treated between 1974 and 2001 (with follow-up to 2003).

The relative risks reported here for vulval cancer after CIN 3, although substantial, are lower than those seen for anal cancer in our previously published study (Coffey *et al*, 2015). We found a RR of 4.03 for incident anal cancer in women with a history of CIN 3 registration before recruitment compared with the RR of 2.68 seen here for vulval cancer. Despite a strong association, vulval cancer remains an uncommon occurrence in women over 50 years with prior cervical neoplasia. In our cohort, 20 cases of vulval cancer arose in the 12 533 women with a history of CIN 3. In other words, ~ 1 in 627 women with a history of CIN 3 have developed vulval cancer, and 2% of the vulval cancers in our study have occurred in women with prior CIN 3 (18 squamous cell carcinomas and 2 melanomas).

A relationship between vulval cancer and obesity has been suspected since the first case series on vulval cancer were published (Green *et al*, 1958; Franklin and Rutledge, 1972). Although one uncontrolled study from Johns Hopkins Hospital (Japaze *et al*, 1977) reported an increased prevalence of obesity among women with vulval cancer, other small case-control studies have found either a non-significant (Newcomb *et al*, 1984; Parazzini *et al*, 1993) or no increase in risk of vulval cancer associated with increased weight or BMI (Brinton *et al*, 1990). Some (Franklin and Rutledge, 1972; Brinton *et al*, 1990; Parazzini *et al*, 1993) have also suggested a relationship between vulval cancer and diabetes that we did not confirm with prospective evidence in our cohort.

Early age at menopause has also been linked with risk of vulval cancer (Green *et al*, 1958; Franklin and Rutledge, 1972). Franklin and Rutledge (1972) reported that 19% of cases had undergone menopause (natural or surgical) before the age of 41 years, and 54% by the age of 45 years. Other small studies that examined the association with age at menopause and related factors (Newcomb *et al*, 1984; Parazzini *et al*, 1993) reported no significant effects of menopausal status, age at menopause, or HT use on risk of vulval cancer. An American case-control study (Brinton *et al*, 1990), which included 209 vulval cancer cases and 348 age-, race-, and residence-matched controls, found an increased risk among women with early ages at natural or surgical menopause: RR 1.89 (95% CI 0.9–3.8) for surgical menopause at <40 years and RR 1.75 (95% CI 0.8–4.1) for natural menopause at <45 years, but this association did not reach statistical significance.

A Danish case-control study (Madsen *et al*, 2008) that included 116 vulval cancer cases found that around half of the vulval tumour samples they examined (compared with 89% of vaginal tumours in

the same study) were positive for HrHPV. When they examined associations between vulval cancer and various exposures stratified by HrHPV positivity, they found significant associations with smoking and cervical neoplasia for only the HrHPV-positive cases.

We found no association between smoking and overall risk of vulval cancer, but when we stratified by tumour histological subtype, we saw a reduced risk of melanocytic vulval cancers among women who reported ever having smoked. This is based on a small number of cases in a subgroup analysis, and therefore must be interpreted with caution. However, other groups have reported similar findings for cutaneous malignant melanoma, with a decrease in incidence and mortality seen in smokers in a recent meta-analysis (Li *et al*, 2015).

A major strength of this study is that it is the first large prospective study to examine the risk of vulval cancer in older women, and our cohort has accrued a relatively large number of incident cases of this relatively uncommon cancer. In addition, CIN 3 status is obtained by routine reporting to cancer registries rather than potentially subjective self-reporting.

Although the number of incident cases of vulval cancers in this cohort is large, the relatively small numbers of cases within exposure subgroups limits our ability to analyse associations overall and by histological subtype. Another weakness lies in our inability to look at vulval dermatosis as an exposure, as we did not ask participants about vulval skin conditions. The vulval cancers that predominate at older ages are likely to be related to chronic vulval conditions, such as lichen sclerosis. It has been reported that 5% of women with symptomatic vulval lichen sclerosis have been found to develop squamous cell carcinomas, although 22% of women presenting with vulval cancers have been found to have 'clinically silent' lichen sclerosis adjacent to their tumours. An interval of ~ 20 years between mean age at diagnosis of vulval lichen sclerosis (54 years) and development of squamous cell carcinoma (74 years) has been reported (Carlson *et al*, 1998), and this is close to the age at which we see a sharp rise in incidence of vulval tumours in the United Kingdom (Figure 1).

We were also unable to distinguish between different types of squamous cell carcinomas. Keratinising squamous cell carcinomas are reported to be the most common histological subtype of vulval cancer in older women, and are less commonly HPV related (de Sanjosé *et al*, 2013). Of the 609 squamous cell carcinomas identified in our cohort, 131 (22%) were specified as keratinising, whereas 447 (73%) were 'squamous cell not otherwise specified'. We only had 8 registrations of tumours with warty (verrucous) morphology and no basaloid squamous cell tumours, the two morphologies most commonly associated with HrHPV infection. We were not able to tell whether tumours with typically HrHPV-related morphologies are truly uncommon in our cohort or whether there is simply a lack of histopathological information captured during tumour registration. At best, tumour histology is a proxy marker of likely HPV association, but it would be helpful to have more robust evidence of histological subtype within the squamous cell group.

CONCLUSION

In a large cohort of over a million women with prospectively collected information on exposures, and just under 14 years of follow-up, we found that history of CIN 3 registration before recruitment, overweight and obesity, and younger age at menopause were associated with an increased risk of incident vulval cancer after the age of 50 years. We also observed a reduction in risk of melanocytic vulval cancers in women who had ever smoked; however, we interpret this finding with caution because of the potential role of chance.

ACKNOWLEDGEMENTS

We thank the women and screening centre staff who participated in the Million Women Study. The Million Women Study is funded by Cancer Research UK (Grant No. C570/A16491) and the UK Medical Research Council (Grant No. MR/K02700X/1). K Coffey is supported by Cancer Research UK Grant Number C38302/A12981 through an Oxford Cancer Research Centre Prize DPhil Studentship. K Gaitskell is supported by Cancer Research UK (CRUK) Grant Number C38302/A17318 through a CRUK Oxford Centre Clinical Research Training Fellowship. K Canfell receives salary support from the National Health and Medical Research Council Australia (CDF # 1082989).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

MILLION WOMEN STUDY COLLABORATORS

Million Women Study Coordinating Centre staff: Hayley Abbiss, Simon Abbott, Rupert Alison, Miranda Armstrong, Krys Baker, Angela Balkwill, Emily Banks, Isobel Barnes, Valerie Beral, Judith Black, Roger Blanks, Kathryn Bradbury, Anna Brown, Benjamin Cairns, Dexter Canoy, Andrew Chadwick, Dave Ewart, Sarah Ewart, Lee Fletcher, Sarah Floud, Alicia Heath, Toral Gathani, Laura Gerrard, Adrian Goodill, Jane Green, Lynden Guiver, Michal Hozak, Isobel Lingard, Sau Wan Kan, Nicky Langston, Kath Moser, Kirstin Pirie, Gillian Reeves, Keith Shaw, Emma Sherman, Helena Strange, Sian Sweetland, Sarah Tipper, Lyndsey Trickett, Clare Wotton, Lucy Wright, Owen Yang, Heather Young.

Million Women Study Advisory Committee: Emily Banks, Valerie Beral, Lucy Carpenter, Carol Dezateux, Jane Green, Julietta Patnick, Richard Peto, Cathie Sudlow.

The NHS Breast Screening Centres that took part in the recruitment of participants were as follows: Avon, Aylesbury, Barnsley, Basingstoke, Bedfordshire and Hertfordshire, Cambridge and Huntingdon, Chelmsford and Colchester, Chester, Cornwall, Crewe, Cumbria, Doncaster, Dorset, East Berkshire, East Cheshire, East Devon, East of Scotland, East Suffolk, East Sussex, Gateshead, Gloucestershire, Great Yarmouth, Hereford and Worcester, Kent, Kings Lynn, Leicestershire, Liverpool, Manchester, Milton Keynes, Newcastle, North Birmingham, North East Scotland, North Lancashire, North Middlesex, North Nottingham, North of Scotland, North Tees, North Yorkshire, Nottingham, Oxford, Portsmouth, Rotherham, Sheffield, Shropshire, Somerset, South Birmingham, South East Scotland, South East Staffordshire, South Derbyshire, South Essex, South Lancashire, South West Scotland, Surrey, Warrington Halton St Helens and Knowsley, Warwickshire Solihull and Coventry, West Berkshire, West Devon, West London, West Suffolk, West Sussex, Wiltshire, Winchester, Wirral, Wycombe.

REFERENCES

Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, Bull D, Elliman J, Harris CL (2001) Agreement between general practice prescription data and self-reported use of hormone replacement therapy and treatment for various illnesses. *J Epidemiol Biostat* 6(4): 357–363.

- Barlow EL, Kang YJ, Hacker NF, Canfell K (2015) Changing Trends in Vulvar Cancer Incidence and Mortality Rates in Australia Since 1982. *Int J Gynecol Cancer* 25(9): 1683–1689.
- Beral V. on behalf of The Million Women Study Collaborative Group (2003) Breast cancer and hormone- replacement therapy: the Million Women Study. *Lancet* 362(9392): 419–427.
- Bobrow KL, Quigley MA, Green J, Reeves GK, Beral V (2012) Persistent effects of women's parity and breastfeeding patterns on their body mass index: results from the Million Women Study. *Int J Obes* 37(5): 712–717.
- Brinton LA, Nasca PC, Mallin K, Baptiste MS, Wilbanks GD, Richart RM (1990) Case-control study of cancer of the vulva. *Obstet Gynecol* 75(5): 859–866.
- Carlson JA, Ambros R, Malfetano J, Ross J, Grabowski R, Lamb P, Figge H, Mihm MC (1998) Vulvar lichen sclerosis and squamous cell carcinoma: a cohort, case control, and investigational study with historical perspective; implications for chronic inflammation and sclerosis in the development of neoplasia. *Hum Pathol* 29(9): 932–947.
- Chaturvedi AK (2010) Beyond cervical cancer: burden of other HPV-related cancers among men and women. *J Adolesc Health* 46(4): S20–S26.
- Coffey K, Beral V, Green J, Reeves G, Barnes I (2015) Lifestyle and reproductive risk factors associated with anal cancer in women aged over 50 years. *Br J Cancer* 112(9): 1568–1574.
- Crawford R, Grignon A-L, Kitson S, Winder DM, Ball SLR, Vaughan K, Stanley MA, Sterling JC, Goon PKC (2011) High prevalence of HPV in non-cervical sites of women with abnormal cervical cytology. *BMC Cancer* 11(1): 473–478.
- de Sanjosé S, Alemany L, Ordi J, Tous S, Alejo M, Bigby SM, Joura EA, Maldonado, Paula Laco J, Bravo IG, Vidal A, Guimerà N, Cross P, Wain GV, Petry KU, Mariani L, Bergeron C, Mandys V, Sica AR, Félix A, Usubutun A, Seoud M, Hernández-Suárez G, Nowakowski AM, Wilson G, Dalstein V, Hampl M, Kasamatsu ES, Lombardi LE, Tinoco L, Alvarado-Cabrero I, Perrotta M, Bhatla N, Agorastos T, Lynch CF, Goodman MT, Shin H-RR, Viarheichyk H, Jach R, Cruz MOLE, Velasco J, Molina C, Bornstein J, Ferrera A, Domingo EJ, Chou C-YY, Banjo A, Castellagué X, Pawlita M, Lloerbas B, Quint WGV, Muñoz N, Bosch FX (2013) Worldwide human papillomavirus genotype attribution in over 2000 cases of intraepithelial and invasive lesions of the vulva. *Eur J Cancer* 49(16): 3450–3461.
- Franklin EW, Rutledge FD (1972) Epidemiology of epidermoid carcinoma of the vulva. *Obstet Gynecol* 39(2): 165–172.
- Green TH, Ulfelder H, Meigs JV (1958) Epidermoid carcinoma of the vulva; an analysis of 238 cases. I. Etiology and diagnosis. *Am J Obstet Gynecol* 75(4): 834–847.
- Japaze H, Garcia-Bunuel R, Woodruff JD (1977) Primary vulvar neoplasia: a review of in situ and invasive carcinoma, 1935–1972. *Obstet Gynecol* 49(4): 404–411.
- Kalliala I, Anttila A, Pukkala E, Nieminen P (2005) Risk of cervical and other cancers after treatment of cervical intraepithelial neoplasia: retrospective cohort study. *Br Med J* 331: 1183–1185.
- Kurman RJ, Carcangiu ML, Herrington CS, Young RH (eds) (2014) *WHO Classification of Tumours of Female Reproductive Organs*, 4th edn. IARC: Lyon, France.
- Li Z, Wang Z, Yu Y, Zhang H, Chen L (2015) Smoking is inversely related to cutaneous malignant melanoma. *Br J Dermatol* 173(6): 1540–1543.
- Madsen BS, Jensen HL, van den Brule AJC, Wohlfahrt J, Frisch M (2008) Risk factors for invasive squamous cell carcinoma of the vulva and vagina—population-based case-control study in Denmark. *Int J Cancer* 122(12): 2827–2834.
- Newcomb PA, Weiss NS, Daling JR (1984) Incidence of vulvar carcinoma in relation to menstrual, reproductive, and medical factors. *J Natl Cancer Inst* 73(2): 391–396.
- Office for National Statistics (2014) Cancer Statistics Registrations, England, Series MB1. Available at: <http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcn:77-262496>.
- Parazzini F, La Vecchia C, Garsia S, Negri E, Sideri M, Rognoni MT, Origoni M (1993) Determinants of invasive vulvar cancer risk: an Italian case-control study. *Gynecol Oncol* 48: 50–55.
- StataCorp (2015) *Stata Statistical Software: Release 14*. StataCorp LP: College Station, TX, USA.
- Townsend P (1987) Deprivation. *J Soc Policy* 16: 125–146.

- van de Nieuwenhof HP, van Kempen LCLT, de Hullu JA, Bekkers RLMM, Bulten J, Melchers WJGG, Massuger LFAG (2009) The etiologic role of HPV in vulvar squamous cell carcinoma fine tuned. *Cancer Epidemiol Biomark Prev* **18**(7): 2061–2067.
- Webb PM (2013) Obesity and gynecologic cancer etiology and survival. *Am Soc Clin Oncol Educ Book*; doi:10.1200/EdBook_AM.2013.33.e222.
- World Health Organization (2000) *International Classification of Diseases for Oncology*, 3rd edn. World Health Organization. Available at: <http://www.who.int/classifications/icd/adaptations/oncology/en/>.
- World Health Organization (2011) *International Statistical Classification of Diseases and Related Health Problems–10th Revision*.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)

**D.3 PREMALIGNANT SOX2 OVEREXPRESSION IN THE FALLOPIAN TUBES
OF OVARIAN CANCER PATIENTS: DISCOVERY AND VALIDATION
STUDIES (EBioMEDICINE, 2016)**

This paper resulted from a project I was involved in, led by Professor Ahmed, and I am listed as a co-author. The project initially aimed to detect early driver mutations in ovarian cancer. One of the genes of interest identified was *SOX2*, and so subsequent experiments further explored its expression and mutation. I was a key contributor to experiments investigating SOX2 expression in the fallopian tubes of women with high-grade serous ovarian cancer compared to women who had had their fallopian tubes removed for other reasons. I performed power calculations, identified suitable cases, scanned stained slides, and conducted image analysis. I also read and commented on drafts of the paper. This work led to the image analysis validation study described in Chapter 8.



Research Paper

Premalignant SOX2 overexpression in the fallopian tubes of ovarian cancer patients: Discovery and validation studies



Karin Hellner^{a,b,1}, Fabrizio Miranda^{a,b,1}, Donatien Fotso Chedom^{a,b,o,1}, Sandra Herrero-Gonzalez^{a,b,1}, Daniel M. Hayden^{c,1}, Rick Tearle^{c,1}, Mara Artibani^{a,b,d}, Mohammad KaramiNejadRanjbar^{a,b}, Ruth Williams^d, Kezia Gaitskell^{a,b}, Samar Elorbany^{a,b}, Ruoyan Xu^{a,b}, Alex Laios^{a,b}, Petronela Buiga^{a,b}, Karim Ahmed^e, Sunanda Dhar^f, Rebecca Yu Zhang^c, Leticia Campo^g, Kevin A. Myers^g, María Lozano^h, María Ruiz-Miróⁱ, Sónia Gatiús^j, Alba Mota^{k,l}, Gema Moreno-Bueno^{k,l}, Xavier Matias-Guiu^j, Javier Benítez^{m,n}, Lorna Witty^o, Gil McVean^o, Simon Leedham^o, Ian Tomlinson^o, Radoje Drmanac^{c,p}, Jean-Baptiste Cazier^{g,q}, Robert Klein^c, Kevin Dunne^c, Robert C. Bast Jr^r, Stephen H. Kennedy^b, Bassim Hassan^s, Stefano Lise^o, María José García^{m,n}, Brock A. Peters^{c,p}, Christopher Yau^{o,t}, Tatjana Sauka-Spengler^d, Ahmed Ashour Ahmed^{a,b,*}

^a Ovarian Cancer Cell Laboratory, Weatherall Institute of Molecular Medicine, University of Oxford, Headington, Oxford OX3 9DS, UK

^b Nuffield Department of Obstetrics and Gynaecology, University of Oxford, Women's Centre, John Radcliffe Hospital, Oxford OX3 9DU, UK

^c Complete Genomics, Inc., 2071 Stierlin Ct., Mountain View, CA 94043, United States

^d University of Oxford, Gene Regulatory Networks in Development and Disease Laboratory, Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, Oxford OX3 9DS, UK

^e Trinity College, University of Cambridge, Cambridge CB2 1TQ, UK

^f Department of Histopathology, Oxford University Hospitals, Oxford OX3 9DU, UK

^g Department of Oncology, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford OX3 7DQ, UK

^h Histopathology Core Unit, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

ⁱ Biobank, Institut de Recerca Biomèdica Lleida (IRBLLEIDA), Lleida, Spain

^j Department of Pathology and Molecular Genetics, Hospital Universitari Arnau de Vilanova, University of Lleida, IRBLLEIDA, Lleida, Spain

^k Department of Biochemistry, Universidad Autónoma de Madrid (UAM), Instituto de Investigaciones Biomédicas "Alberto Sols" (CSIC-UAM), IdiPAZ, Madrid, Spain

^l MD Anderson international Foundation, Madrid, Spain

^m Human Genetics Group, Spanish National Cancer Research Center (CNIO), Madrid, Spain

ⁿ Biomedical Network Research Centre on Rare Diseases (CIBERER), Spain

^o Wellcome Trust Centre for Human Genetics, NIHR Biomedical Research Centre, Roosevelt Drive, Oxford OX3 7BN, UK

^p BGI-Shenzhen, Shenzhen, China

^q Centre for Computational Biology, University of Birmingham, Edgbaston B15 2TT, UK

^r Department of Experimental Therapeutics, M.D. Anderson Cancer Center, University of Texas, USA

^s Tumor Growth Control Group, Sir William Dunn School of Pathology, South Parks Road, Oxford OX1 3RE, UK

^t Department of Statistics, 1 South Parks Road, Oxford OX1 3TG, UK

ARTICLE INFO

Article history:

Received 31 May 2016

Received in revised form 30 June 2016

Accepted 30 June 2016

Available online 2 July 2016

Keywords:

Ovarian cancer

Fallopian tube

BRCA mutations

SOX2

Screening

Precancer

ABSTRACT

Current screening methods for ovarian cancer can only detect advanced disease. Earlier detection has proved difficult because the molecular precursors involved in the natural history of the disease are unknown. To identify early driver mutations in ovarian cancer cells, we used dense whole genome sequencing of micrometastases and microscopical residual disease collected at three time points over three years from a single patient during treatment for high-grade serous ovarian cancer (HGSOC). The functional and clinical significance of the identified mutations was examined using a combination of population-based whole genome sequencing, targeted deep sequencing, multi-center analysis of protein expression, loss of function experiments in an in-vivo reporter assay and mammalian models, and gain of function experiments in primary cultured fallopian tube epithelial (FTE) cells. We identified frequent mutations involving a 40 kb distal repressor region for the key stem cell differentiation gene SOX2. In the apparently normal FTE, the region was also mutated. This was associated with a profound increase in SOX2 expression ($p < 2^{-16}$), which was not found in patients without cancer ($n = 108$). Importantly, we show that SOX2 overexpression in FTE is nearly ubiquitous in patients with HGSOCs ($n = 100$), and common in BRCA1-BRCA2 mutation carriers ($n = 71$) who underwent prophylactic salpingo-oophorectomy. We propose

* Corresponding author at: Ovarian Cancer Cell Laboratory, Weatherall Institute of Molecular Medicine, University of Oxford, Headington, Oxford OX3 9DS, UK.

E-mail address: ahmed.ahmed@obs-gyn.ox.ac.uk (A.A. Ahmed).

¹ These authors contributed equally to this work.

that the finding of SOX2 overexpression in FTE could be exploited to develop biomarkers for detecting disease at a premalignant stage, which would reduce mortality from this devastating disease.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Because of late presentation and chemotherapy resistance ovarian cancer is the deadliest amongst all gynecological malignancies. Over 80% of ovarian tumors are high-grade serous ovarian cancers (HGSOCs), which is a particularly fatal type due to its initial asymptomatic but aggressive growth behavior. Current methods of detection have been successfully implemented for detection and possible reduction of mortality from ovarian cancer (Menon et al., 2015; Drescher et al., 2013). However, such methods are only capable of the detection of established invasive cancers. Understanding the natural history of the disease and the discovery of novel markers for detection at a premalignant stage will enable the effective control of ovarian cancer.

In this work, we prospectively analyzed the genomic composition of a single tumor over a three-year period to identify driver mutations that may have contributed to the initiation of the tumor. We identified non-coding mutations that cluster near genes involved in stem cell regulation. We established that one mutation is located in a previously unrecognized repressor element of *SOX2*, an important stem cell gene, and is associated with induction of *SOX2* expression. We demonstrate that the expansion of *SOX2*-expressing cells within the fallopian tube epithelium is a common feature of HGSOCs, a crucial finding that opens new avenues for early disease detection prior to clinical presentation.

2. Materials and Methods

2.1. Overall Description of the Study Design

The clinical samples for this translational study were obtained from patients recruited to the Gynecological Oncology Targeted Therapy Study 01 (GO-Target-01) and the Oxford Ovarian Cancer Predict Chemotherapy Response Trial (OXO-PCR-01) under research ethics approval number 11-SC-0014 and 12-SC-0404, respectively. We performed intraoperative video recording to document sampling sites (Supplementary video). Strict standard operating procedures were used to diminish the risk of DNA cross-contamination during sample collection and processing. Whole genome sequencing (WGS) of laser capture microdissected tumor islets ($n = 30$, Supplementary Fig. S1A) and bulk tissue samples of a single HGSOC (patient study ID: 11152). We obtained WGS data from 39 samples from three independent data sets and a tumor recurrence set (Supplementary Fig. S1B). Samples were obtained from different locations before chemotherapy, after neoadjuvant chemotherapy and approximately two and half years later at the time of first recurrence (Fig. 1, Supplementary Table 1). Complete macroscopic clearance at all sites (Supplementary video) as well as microscopic clearance of the peritoneal implants at sites A and B were documented following chemotherapy. Microscopic residual chemoresistant disease

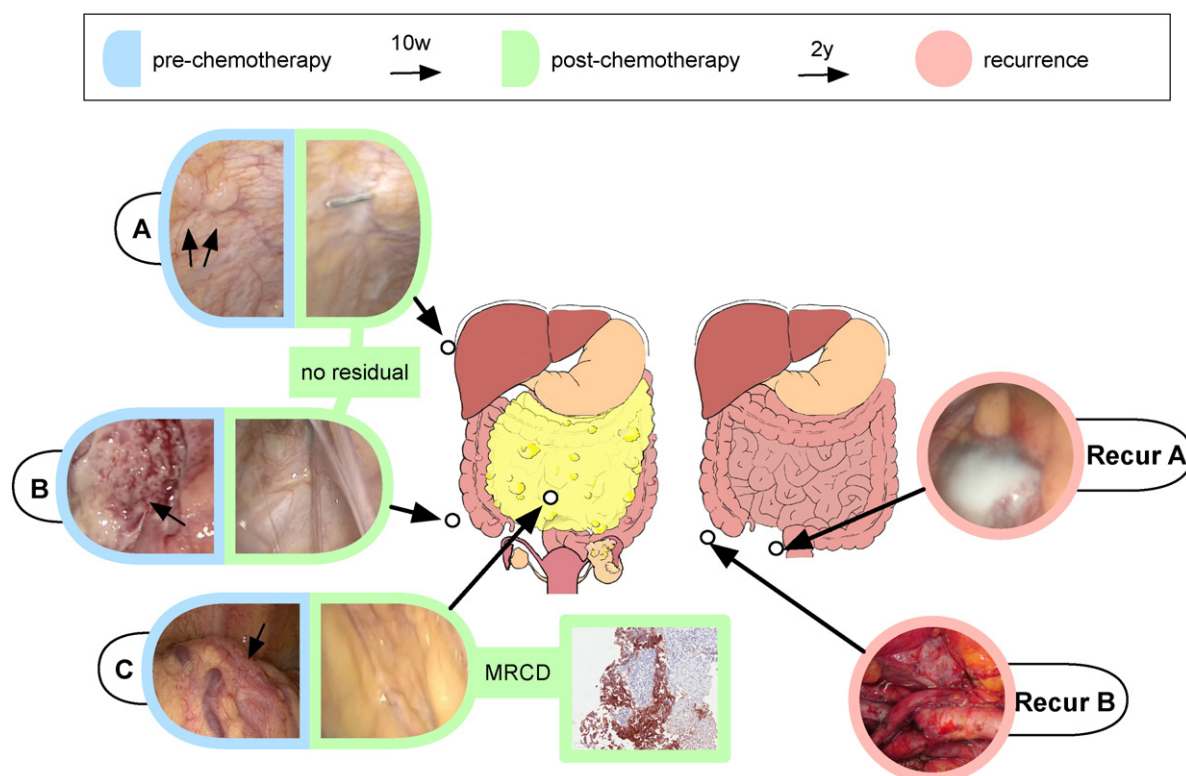


Fig. 1. An ovarian cancer model for investigating primary chemotherapy resistance. A diagram illustrating the sites from which the biopsies were obtained in patient 11152 and the corresponding intra-operative images of the biopsy sites. The sub-diaphragmatic peritoneum (site A), the para-cecal peritoneum (site B) and the omentum (site C) were sampled in the primary tumor. A para-rectal mass (Recur A) and a pelvic node (Recur B) were sampled at presentation of disease recurrence. Note the complete macroscopic resolution of the tumor following chemotherapy (also see Supplementary video). TP53 immunohistochemical staining of a tumor islet from MRCD is also presented.

(MRCD) was only detected at the omentum in position C. Sequencing of tumor islets was performed using long fragment read (LFR) WGS (Peters et al., 2012) of 30 tumor islets (typically comprising only tens of cancer cells). A table of variants was generated and filtered using “cgatools” and custom scripts. GREAT software (McLean et al., 2010) was used for gene ontology analysis. Targeted sequencing was carried out for validation of mutations using 100 bp paired-end sequencing and read mapping was performed using STAMPY software (Lunter and Goodson, 2011), whilst variants were identified using PLATYPUS (Rimmer et al., 2014). Digital PCR was performed using the Droplet Reader (BioRad) and analyzed with the QuantaLife software (BioRad). For the in-vivo studies, fertilized chicken eggs were electroporated with enhancer constructs and analyzed as previously described (Betancur et al., 2010; Simoes-Costa et al., 2012). Targeted deletion of the BB5 region was performed using Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) (Cong et al., 2013). Chromatin immunoprecipitation (ChIP) assays were carried out in primary fallopian tube epithelial cells (Cong et al., 2013) and a fraction of the products was used for either real-time PCR reactions or high throughput sequencing. Immunohistochemistry studies were carried out on the Leica Bond Max staining system including the appropriate positive and negative controls. Detailed materials and methods are available in the Supplementary methods section.

2.2. Translational Studies

2.2.1. Ethical Approval

Tumor islets and blood samples for whole genome sequencing were obtained from case 11152 who provided written consent to participate in the prospective biomarker validation study Gynecological Oncology Targeted Therapy Study 01 (GO-Target-01) under research ethics approval number 11/SC/0014. Targeted sequencing was performed on samples from patients who participated in the same study and patients who participated in the prospective Oxford Ovarian Cancer Predict Chemotherapy Response Trial (OXO-PCR-01), under research ethics approval number 12/SC/0404. Necessary informed consents from study participants were obtained as appropriate.

2.2.2. Clinical History and Sample Collection of Patient 11152

Patient 11152 presented with radiological evidence of at least stage IIIC ovarian cancer with evidence of multiple omental and peritoneal nodules, subcapsular splenic nodules and left paracardiac lymph node enlargement and elevated CA125 tumor marker. The patient had a diagnostic laparoscopy and biopsy to confirm the diagnosis of HGSOc. On the day of the procedure, tissue samples were retrieved, split in halves and either saved in cryovials and immersed in dry ice within minutes from obtaining the biopsy in theatre or formalin fixed for standard histological diagnosis. Research samples were stored in -80°C . In addition, a blood sample was obtained in theatre and saved in sterile EDTA-collection tubes. The patient received three cycles of paclitaxel and carboplatin and had near complete response. Following ten weeks from the first laparoscopy, the patient underwent a diagnostic laparoscopy to evaluate chemotherapy response, prior to proceeding, in the same session, to standard debulking surgery. Guided by the first video-laparoscopy, samples were obtained from the same sites from which the initial biopsies were taken.

2.2.3. Precautions for Tissue Handling to Diminish the Risk of Cross Contamination of DNA

Samples were collected from -80°C to a box of dry ice and transported within the same building to the cryostat room. Individual cryovials were obtained from the dry-ice box and immediately placed within a pre-cooled cryostat stage at -25°C . The cryostat (CryoStar 70X, Thermo Scientific) was not previously used for routine diagnostic or research work. The cryostat stage and dissection board were cleaned by ethanol prior to use for individual samples, and gloves were changed

before cutting each new tissue biopsy. For each sample, a single dissection blade (MB DynaSharp Microtome Blade, Thermo Scientific) and single sterile forceps (that was cleaned, individually wrapped and autoclaved prior to use) were used. Samples were obtained from the cryovials and placed on a sterile dish and cut using disposal single-use blades to obtain a piece of tissue for processing. These tissue pieces were then placed on individual new dissection discs for snap freezing in optimal cutting temperature (OCT) compound (NEG-50, Richard-Allan Scientific) prior to microtome cutting.

2.2.4. Section Processing for Laser Capture Microdissection

The first tissue section was mounted onto regular glass slides (SuperFrost Plus, VWR International) for hematoxylin (Hematoxylin solution, Gill No. 3, Sigma) and eosin (Eosin Y solution, Sigma) staining (H&E), according to manufacturer's instructions, followed by six to ten sequential tissue sections at $6\text{ }\mu\text{m}$ thickness onto polyethylene naphthalate membrane (PEN) glass slides (MembraneSlide 1.0 PEN, Zeiss) that were pre-exposed to UV light for 30 min. Slides were immediately stored at -80°C . Nuclease-free technique was used throughout the procedure and buffers and alcohol solutions were cooled to 4°C and used fresh each time. Each H&E slide was reviewed by a gynecological oncology pathologist (SD) to confirm the presence of cancer cells and delineate their location. PEN slides were dipped in 50% ethanol for fixation and rinsed in H_2O to remove excess OCT compound. The slides were stained with cresyl violet (Sigma Aldrich) at a concentration of 0.1% (weight/volume) in 50% ethanol for 15 s, rinsed in 50% ethanol and immediately used for microdissection. For formalin-fixed, paraffin embedded (FFPE) material, $6\text{ }\mu\text{m}$ sections were cut onto activated PEN slides and dried at 56°C overnight, then dewaxed in Xylene (Sigma) and rehydrated through graded alcohols to water, then briefly dipped in 1% methyl green (Sigma), washed in H_2O . The slides were dried at 37°C for 1 h and then used for microdissection. Laser capture microdissection was performed on a PALM Laser Microdissection System (Zeiss) and the cut tissue was catapulted into 200 μl membrane caps (AdhesiveCap 200 opaque, Zeiss, Jena, Germany) and immediately stored on dry ice. Images of empty caps as well as target area in $5\times$, $10\times$ and $20\times$ magnification were obtained prior to microdissection for documentation. To maintain DNA integrity, slides were kept on dry ice until microdissection and microdissection was performed for no longer than 45 min per slide. Caps were stored at -80°C until sequencing.

2.2.5. Macrodissection of FTE

Serial unstained FFPE tissue sections ($10\text{ }\mu\text{m}$) were aligned with the corresponding H&E stained section in which the location of the FTE was pen marked. With a sterile, single-use scalpel blade tip the FTE tissue was scrapped off the glass slide and mounted onto 200 μl a membrane cap (AdhesiveCap 200 opaque, Zeiss, Jena, Germany). To avoid cross contamination, the slides and caps were placed in sterile tissue culture dishes. DNA extraction and amplification was performed using the REPLig FFPE kit (Qiagen). For macrodissection of the p53 signature from FFPE immuno-stained sections the slides were incubated overnight in Xylene at 37°C . The coverslip was lifted and the slide soaked in Xylene for 30 min, then washed in gradient ethanol. The tissue was macrodissected using a scalpel blade tip. DNA was extracted with the Arcturus® PicoPure® DNA isolation kit (Life technologies) and amplified using the REPLig mini kit (Qiagen) according to manufacturer's instructions. DNA was quantified as described below.

2.3. DNA Extraction

For DNA extraction from frozen tissues 10 to 20 scrolls of tissue were cut at $60\text{ }\mu\text{m}$ thickness using a cryostat as described above. For FFPE samples, 10 to 20 scrolls of $20\text{ }\mu\text{m}$ thick sections were dewaxed with Xylene and washed with 100% ethanol. DNA was extracted using the DNeasy blood and tissue kit (Qiagen) according to manufacturers instructions. DNA concentrations were quantified using the PicoGreen

(Quant-iT™ PicoGreen® dsDNA Assay Kit, Life technologies), A260/280 ratio and absorption spectra were generated using a spectrophotometer (Nanodrop ND-1000, Thermo Scientific) and broad range Qubit system (Invitrogen) and quality was checked using a 1% agarose E-gel (Invitrogen).

2.4. Sequencing

2.4.1. Targeted Sequencing

LFR sequencing and whole genome sequencing were performed as previously described (Dean et al., 2002; Drmanac et al., 2010). A summary of the clinical characteristics of patients for whom targeted sequencing was performed is presented in Supplementary Table 2. For sequencing the 2 Mb flanking the *SOX2* gene, the online NimbleDesign tool (<http://www.nimblegen.com/products/nimbledesign/index.html>) was used to design capture primers (Roche). For sequencing of the test set the TruSeq® DNA HT Sample Preparation Kit (Illumina) was used to allow for multiplexing and captured using the Seq Cap EZ Choice Library (Roche). Both kits were used according to manufacturer's instructions. Library quality control was carried out using the broad range Qubit system (Invitrogen) and the 2200 TapeStation (Agilent). Sequencing was carried out on a HiSeq2500 (Illumina) using TruSeq® Rapid SBS 100 bp paired end sequencing. For sequencing the validation set of matched tumor/normal DNA pairs, a more focused library of capture primers was designed targeting ~1.6 Mb flanking *SOX2* (chr3: 180,806,202–182,429,714). DNA from paired samples and ChIP products were fragmented (S2, Covaris) and sample libraries were constructed and multiplexed using Seq Cap EZ kit (Roche) according to manufacturer's instructions. Capture was performed as described above. Quality control was carried out on a 2100 Bioanalyzer (Agilent) and library concentrations were measured with the high sensitivity Qubit system (Invitrogen). Sequencing was performed on a MiSeq platform (Illumina) using v3 MiSeq chemistry. To increase coverage, the libraries for the blood sample and the microdissected tubal epithelium of case 11152 were recaptured and sequenced using v2 MiSeq chemistry, 100 bp paired end.

2.4.2. Sanger Sequencing

For DNA sequencing using dye-terminator method, *TP53* exon 8 was amplified and sequenced using primers *TP53*-forward 'GGGTGCAGTTA TGCCTCAGATT' and *TP53*-reverse 'CGGCATTTTGAGTGTAGACTGG' as previously described (Ahmed et al., 2010). *SOX2* BB5 was amplified and sequenced using the BB5-forward 'CACCATGTGAATCATCTCG' and BB5-reverse 'ACCAGGTGTCCGAGAGTACG' primers. PCR was performed using the high fidelity DNA Phusion polymerase (NEB) as per manufacturer's instructions. Sequencing was performed for the rare variants identified in patients (Supplementary Table 3) using the primers listed in Supplementary Table 4.

2.4.3. Digital Droplet PCR

Digital droplet polymerase chain reaction (PCR) was performed on duplicate samples. Primers 5833217_F; 5'-ACCTACTAGACCCCA GGCAAG-3' and 5833217_R; 5'-GGCGCAGGAGGAGACC-3' were used to amplify a 60 bp amplicon containing the BB5 nucleotide and either detected using 5833217_V; 5'-CCTGGGACCCAAACC-3' VIC-labeled probe for wild type or 5833217_M; 5'-CTGGCACCACAAACC-3' FAM-labeled probe for mutant amplicons (TaqMan® SNP Genotyping Assays, custom design, Roche Molecular Systems). *TP53* mutation was quantified using primers 22410689_F; 5'-CTGTGCGCCGGTCTCT-3' and 22410689_R; 5'-TGGGACGGAACAGCTTTGAG-3' to amplify a 64 bp amplicon and detected using 22410689_V 5'-TGCGTGTGTGCTG-3' VIC-tagged probe for wild type and 22410689_M; 5'-TGCGTGTGTGCTG-3' FAM-tagged probe for mutant amplicons. Reactions were prepared using droplet digital PCR Super Mix (BioRad) and standard PCR performed according to manufacturer's instructions. Amplification events were detected with a digital PCR plate reader

(QX100 Droplet Reader, BioRad) and data was analyzed using the QuantaSoft Software (Version 1.3.2.0, BioRad). Average droplet count was 11,728 per sample. Samples with <7000 droplets were excluded from the analysis.

2.5. Sequencing Analysis

2.5.1. Whole Genome Sequencing

Reads were mapped to the reference genome (GRCh37) and variants were called by local de novo assembly as previously described (Carnevali et al., 2012). A table of variants across the various tumor and normal genome assemblies was generated and filtered using the Complete Genomics cgatools program and custom scripts. Sequence data were deposited in the European Genome-Phenome Archive (EGA) with the following accession number: EGAS00001001909.

2.5.2. Performing Gene Ontology Enrichment Analysis

The functional prediction of *cis*-regulatory regions was performed using the Genomic Region Enrichment of Annotation Tool (GREAT) as previously described (McLean et al., 2010). In brief, GREAT assigns regulatory domains for each gene that consists of a basal domain (5 kb upstream and 1 kb downstream of transcription start site [TSS]) plus extension of up to 1 Mb, but not beyond 1 Mb, in both directions to the nearest gene basal domain. The enrichment for a particular biological process gene ontology (which consists of a number of genes) was computed by obtaining the ratio of the fraction of foreground (FG) variants (i.e. the ancestor variants) that mapped to genes of a gene ontology x to the fraction of background (BG) variants (e.g. progeny variants) that mapped to the same gene ontology. To be regarded as a gene ontology hit, we required that a gene ontology had a false discovery rate "Q value" of <0.01 and a number of genes supporting a particular ontology of more than ten. The BG was defined as either the progeny variants (variants present in two or more tumor sites), the whole genome or the number of germline variants as indicated in the text. Permutation analysis was based on the assumption that random samples from the background of equal size to the FG should not give higher enrichment than the one observed for the true FG. To test this, we obtained 10,000 samples of 750 variants from the BG and computed the enrichment for each and counted the number of times in which that enrichment was higher than the one observed by the true FG and presented the result (p -value) as the fraction of the count from the total number of random samples. The number of genes that GREAT assigned to the gene ontology "stem cell differentiation" was 73 and this was the basis of the analysis described above. We repeated the above analysis using the total number of human genes (283) assigned to the same gene ontology "GO:0048863" at the Gene Ontology data base (<http://amigo.geneontology.org/cgi-bin/amigo/go.cgi>) and by using a subset of genes ($n = 81$) that had an experimentally verified link to stem cell differentiation by selecting the human genes that satisfied any of the following terms: IDA; inferred from direct assay, IEP; inferred from expression pattern, IGI; inferred from genetic interactions, IMP; inferred from mutant phenotype.

2.5.3. Analysis of Targeted Sequencing Data

Reads were mapped to the reference genome (GRCh37) using STAMPY software (Lunter and Goodson, 2011) and variants were identified using an in-house developed software, PLATYPUS (Rimmer et al., 2014). Downstream analysis was performed using in-house developed scripts. There were a total of 37,291 variants identified by sequencing. Further analysis focused on single nucleotide substitutions in cancer samples ($n = 18,456$ in 33 samples). To identify high quality rare variants, the analysis was restricted to variants that were present in less than five samples, were not called in the 1000 Genomes Project (Abecasis et al., 2012), and had a high quality score (i.e. flagged as "pass" or "allele bias" by PLATYPUS). Only 861 variants met these criteria and these are shown in Supplementary Table 5. In order to

compute whether the observed rare variants displayed regional enrichment in the ovarian cancer set we adopted a simulation approach. First, we obtained minor allele frequencies (MAF) for common SNPs (defined as having MAF > 5%) from the 1000 Genomes Project in the 2 Mb region flanking SOX2. We then simulated genotypes, based on those minor allele frequencies, for samples consisting of 33 individuals (we ignore linkage effects) and counted the number of heterozygotes or non-reference homozygotes observed across the 33 simulated profiles. This process was repeated 1000 times in order to obtain an average profile of common SNP density across the region. We then examined 40 kb windows, spaced at overlapping intervals of 100 bp, and counted the number of rare variants observed across the 33 ovarian cancers. For each window, we then computed the test statistic (observed number of rare variants in ovarian cancers/expected number of common SNPs). In order to assess whether the calculated ratios were significantly enriched, we randomly sampled 100 subsets of 33 individuals from the wellerly (well elderly individuals) dataset and computed the same statistic for each subset by counting rare variants falling in any particular window. We then computed a *p*-value from the order-statistic of the observed test statistic in the ovarian cancers compared to those of the wellerly data.

In addition, the 861 variants were annotated using data from the Encyclopedia of DNA Elements (ENCODE) (Bernstein et al., 2012) to identify those that were within regulatory regions as reported by digital genomic footprint (Neph et al., 2012) downloaded from (ftp://ftp.ebi.ac.uk/pub/databases/ensembl/encode/integration_data_jan2011/byDataType/footprints/jan2011/), by DNaseI hypersensitivity (Thurman et al., 2012) downloaded from (<http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeRegDnaseClustered/>) or by ChIP-Seq (Yip et al., 2012) experiments downloaded from (<http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeRegTfbsClustered/>). In order to assess the degree of local enrichment of variants we first computed the global rate of variant occurrence over the 2 Mb region (113 variants/2,000,000 bp). We then considered a series of 40 kb overlapping windows spaced at 100 bp intervals spanning the 2 Mb. For each window, we counted the number of variants observed in the tumors and computed the probability of observing (at least) this many variants under the null hypothesis of uniform variant occurrence. We assumed that, under the null hypothesis, the number of observed variants follows a Poisson distribution with a rate parameter given by the global rate over the region.

2.5.4. Analysis of ChIP-sequence Tracks

Following the ENCODE guidelines (Landt et al., 2012), read quality was first assessed using FastQC (v.0.10.1; <http://www.bioinformatics.babraham.ac.uk>). Low quality reads were filtered by using Trim galore (v 0.3.1; <http://www.bioinformatics.babraham.ac.uk>) for paired reads. The 3' and 5' adapters were trimmed using the Cutadapt program implemented in Trim galore. Only reads with a length > 60 bp after trimming were retained. Reads were then mapped against the human reference genome (build hg19) using bowtie (Langmead and Salzberg, 2012). Only non-duplicated aligned reads were retained. Multiple alignment of read was not allowed. To identify enriched regions, peaks were called on the immune-precipitated samples versus their input controls using Model-Based Analysis of ChIP-seq 2 (Feng et al., 2012) (MACS2 version 2.0.10) with a *p*-value threshold of 0.001 and default parameters. Signal track was built following the guideline at <https://github.com/taoliu/MACS/wiki/Build-Signal-Track>. Exploratory analysis and generation of figures was performed in R using custom scripts.

2.5.5. Motif Analysis for the BB5 Variant

The motif UW.Motif.0169 overlapping the BB5 variant was discovered from digital footprints data. The logo plot in Supplementary Fig. S6A was generated using the webLogo3 software (<http://weblogo.threeplusone.com/>).

2.6. Fallopian Tubes Primary Epithelial Cell Culture

Patients scheduled to undergo surgical procedures provided written consent, prior to surgery, agreeing to participate in the study. The infundibular region of the fallopian tube was isolated, dissected and opened to reveal the lumen. Fallopian tubes were incubated in 15 ml conical tubes containing 0.5% trypsin and 0.1% DNaseI in MEM for 1 h at 37 °C, with shaking. The supernatant, containing the epithelial cells, was removed and mixed with 10% FBS in DMEM. Fallopian tube epithelial cells (FTEC) were centrifuged and plated in 10% DMEM. The purity of epithelial cells was checked by immunofluorescence. Optional purification using CD326 microbeads (Miltenyi Biotech) was if further purification was required.

2.7. Viral Transduction

pLx302-SOX2 was generated using Gateway system (Invitrogen) according to the manufacturer's instructions. SOX2 was cloned into pDONOR221 (Invitrogen) from pMXs-Sox2 (Addgene) plasmid by BP Gateway clonase (Invitrogen) to generate entry clones. Using LR Gateway clonase (Invitrogen) SOX2 was cloned into the destination vector pLx302 (a gift from Dr. David Root (Yang et al., 2011); Addgene plasmid 25896). Packaging cells (HEK-293T) were co-transfected with the packaging vector p8.91, the envelope plasmid pMDG (gifts from Dr. Didier Trono, University of Geneva, Switzerland) and pLx302-SOX2 as previously described (Besnier et al., 2002). Following 72 h the HEK293-T medium containing the virus was collected, filtered through a 0.45 µm Minisart NML Syringe Filter (Sartorius) and stored at –80 °C for later transduction.

2.8. Cloning, Mutant Generation and Chicken Embryo Transfection

The 1 kb regions flanking the BB5 SNP was cloned from Human genomic blood DNA by PCR amplification using Phusion high fidelity polymerase (NEB) according to the manufacturer's instructions using the following primers BB5_F; 'TTTTTCGTCTCgccaggTTACTCCAATA TGAGAGATAAGAGCA' and BB5_R; 'TTTTTCGTCTCcaacagCGCTCACACG GTGATTAGAA', Test1_F; 'TTTTTCGTCTCgccaggCATTACTGGCAGCTGA GGGG' and Test1_R; 'TTTTTCGTCTCcaacagTGATTTCCCTGGGCAGACA', Test2_F; 'TTTTTCGTCTCgccaggCTACTAGACCCAGGCAAGG' and Test2_R; 'TTTTTCGTCTCcaacagTGATTTCCCTGGGCAGACA', Test3_F; 'TTTTTCGTCTCgccaggTCCCTGTCTCTACTCTCT' and Test3_R; 'TTTTTCGTCTCcaacagTGATTTCCCTGGGCAGACA'. Products were cloned into a ptk citrine – BsmBI vector. Restriction digestion of the PCR products and the vector was performed using BsmBI enzyme (NEB) and the digested product was inserted into the vector using a T4 ligase (NEB). Point mutants were generated using QuikChange® II Site-Directed Mutagenesis Kit (Stratagene) as per manufacturer's instructions using the citrine vector containing each region and the primer sets SOX2_BB5mut_F; 'ACCTGGGCCTGGCACCAACCCCTT' and SOX2_BB5mut_R 'AAGGGTTTGGGTGCCAGGCCAGGT'. DNA sequences of all cloned PCR products were verified with direct sequencing.

Fertilized chicken (*Gallus gallus*) eggs (Henry Stewart & Co. Ltd, Louth, UK) were incubated at 37 to 38 °C for approximately 20 to 24 h prior to electroporation. In the UK no license is required at the embryo stage used for this work. The entire epiblast of stage four chicken embryos (staged according to Hamburger and Hamilton (Hamburger and Hamilton, 1992)) was electroporated with enhancer constructs, cultured using modified new culture (Sauka-Spengler and Barenbaum, 2008) and analyzed as previously described (Betancur et al., 2010; Simoes-Costa et al., 2012) (*n* = 10 to 15 per group). Neither randomization nor blinding was carried out. Non-vital or damaged embryos were excluded from further analysis. Immunofluorescence for detection of endogenous Sox2 expression and co-localization with citrine expression was carried out as previously described (Betancur et al., 2010; Simoes-Costa et al., 2012) using anti-Sox2 antibody (ab97959, Abcam) or anti-GFP antibody (MAB3580, Millipore) and detected using alexa-488 conjugated and alexa-568 conjugated secondary antibodies

(Invitrogen) and observed using an inverted confocal microscope (LSA 510 META, Zeiss).

2.9. Chromatin Immunoprecipitation (ChIP) Assay

ChIP was performed on FTEC and HEK293T cells (purchased from ATCC, mycoplasma tested) using the commercially available ChIP-IT express enzymatic kit (Active Motif) according to manufacturer's instructions and the following antibodies: H3K27ac (C15410174 lot: A.7071-001P; Diagenode) and EZH2 (C15410039; lot: 003; Diagenode). Briefly, one million cells were cross-linked using formaldehyde 1% for 10 min and the reaction was stopped by adding glycine for 5 min at room temperature. Cells were then washed twice with ice cold PBS containing $1 \times$ PMSF, and then resuspended in cell lysis buffer containing $1 \times$ protease inhibitor cocktail and PMSF and rotated at 4 °C for 30 min. Nuclei were extracted using a Dounce homogenizer (pestle B) and treated with an enzymatic shearing cocktail at 37 °C for 20 min. Digested chromatin was immunoprecipitated at 4 °C overnight by adding protein G magnetic beads and rabbit IgG (cat. N. 12-370; lot. N. 1969696; Millipore) as a negative control. Following washes, reverse cross-linking and proteinase K treatment the DNA was purified using the Chromatin IP DNA purification kit (Active Motif). Sequencing of ChIP material was performed as described above. Real-time PCR was performed in a 96-well plate qPCR machine (ABI 7000, Applied Biosystems). Reactions were set up using SybrGreen (Applied Biosystems) according to the manufacturer's protocols using 1 µl of each ChIP reaction or 1: 100 to 200 dilution of the input fraction. The $\Delta\Delta C_t$ method was used for quantification and calculations performed according to ChIP-qPCR data analysis instructions (Supper-Array, Bioscience Corporation). The following primers for the BB5 region were used: 5833217_F 5'-ACCTACTAGACCCAGGCAAG-3' and 5833217_R 5'-GGCGCAGGAGGAGACC-3'. For EZH2 inhibition, cell were treated for 96 h with 5 µM of UNC1999 (Sigma) before chromatin extraction.

2.10. CRISPR Vector Construction, Cell Culture and Transfection

Vector px330 was used as previously described (Cong et al., 2013). A pair of oligonucleotides (CRISPR-BB5-F; 5'-CACCGAGGGTTTGG GTCCCAGGCC-3' and CRISPR-BB5-R; 5'-AAACGGGCTGGGACCC AAACCTC-3') encompassing the BB5 nucleotide (underlined) and extending up to a protospacer adjacent motif [AGG] (not shown) were annealed, phosphorylated, and ligated to the linearized vector.

HEK293 cells (ATCC) were maintained in DMEM (Invitrogen) supplemented with 10% fetal bovine serum and 100 U/ml penicillin/streptomycin and incubated at 37 °C and 5% CO₂. HEK293 cells were transfected with the construct px330-BB5 using FUGENE HD (Promega) according to the manufacturer's instructions. After two weeks, the DNA for each clone was extracted using DNeasy blood and tissue kit (Qiagen) and the region flanking the BB5 nucleotide was amplified using the following pair of primers: SH-BB5-01-F; 5'-TCCAATATGAGAGATAA GAGCAC-3' and SH-BB5-01-R; 5'-GCTGAAAAGACCAACTTAAAC-3'. HEK293T cells were transfected and two days later single cells were plated in 96-well plates and harvested for DNA collection on the indicated dates. The number of positive clones carrying bi-allelic mutations (nominated as homozygous), the number of clones carrying mono-allelic mutation (nominated as heterozygous) and clones with no deletions (nominated as negative) are reported in Supplementary Table 6.

2.11. Immunofluorescence

FTEC or SKOV3 cells were fixed with 4% (w/v) formaldehyde in PBS for 4 min and then permeabilized overnight with 100% ethanol at -20 °C. Cells were then blocked with 1.5% BSA in TBS for 1 h and incubated with primary mouse antibody against pan-cytokeratin (ab7753, Abcam), primary rabbit antibody against MYC (9402S, Cell Signaling Technology), primary rabbit antibody against SOX2 (3579, Cell

signaling Technology), primary mouse antibody against SOX2 (4900, Cell Signaling Technology) or primary rabbit antibody against PAX8 (10336, Proteintech) for 1 h. Cells were washed using washing buffer (TBS + Triton 0.2% + SDS 0.04%) and then incubated with donkey anti-mouse or donkey anti-rabbit conjugated with Alexa Fluor 488 or Alexa Fluor 546 secondary antibodies (Invitrogen), respectively. After three washes cells were mounted using DAPI-containing mounting media (Vector Laboratories).

2.12. Immunohistochemistry

A summary of patients from both groups (discovery set and validation sets) for whom SOX2 IHC was performed is presented in Supplementary Table 7. Power calculations were performed to determine the minimum specimen number for the validation set based on the data from the discovery set (mean % for benign: 3, s.d.: 4; mean % for HGSOs: 14, s.d.: 10, difference between means: 11%) assuming a desire for 90% power and a 5% significance level. To identify the difference in 3 + staining intensity for the above scenario we estimated a group size of ten cases.

Tissue sections of 2.5 µm or 4 µm thickness were cut from FFPE tumor or control samples. Automated staining was carried out with the Leica Bond Max autostainer (Leica Microsystems) or Autostainer plus Link 48 (Dako®). In short, antigen retrieval at 100 °C for 20 min was followed by primary antibody incubation with the rabbit anti-SOX2 (3579, Cell Signaling Technology) or IgG control for up to 1 h then detection using the BOND™ Polymer Refine Detection System (DS9800, Leica Biosystems) as per manufacturer's instructions. Double staining was achieved using the BOND™ Polymer Refine Detection System and the BOND™ Polymer Refine Red Detection System (DS9390, Leica Biosystems) according to the manufacturer's instructions. Stained slides were scanned at 20× and 40× magnification on the Aperio slide scanner (Aperio) or Mirax scan (Zeiss) fitted with a 40×/0.95 Plan Apochromat objective lens (Zeiss) for the second validation set (CNIO; Spanish National Cancer Research Centre). The ImageScope software (v11.2.0.780, Aperio) was used for quantification of nuclear staining for the discovery set and the first validation set. For scoring of SOX2 positivity in fallopian tubes, only the tubal epithelium was marked by using the "negative pen tool" to exclude stroma. The marked FTE was analyzed with the program algorithm "nuclear v9" which scored the staining of all nuclei within the marked area and assigned scores of 0 for not detectable signal, +1 for weak staining, +2 for moderate staining and +3 for strong staining. Nuclear positivity was confirmed on selected areas using the "deconvolution" algorithm to subtract SOX2 from underlying hematoxylin staining. The diagnosis and clinical details are provided in Supplementary Table 7. For scoring of SOX2 expression in HGSOs tumor foci were marked and benign tissue excluded and the procedure completed as described above. For scoring the second validation set (CNIO), the AxioVision 4.6.3 software (Zeiss) pattern recognition training software was used. Algorithm parameters were set by using the AxioVision Wizard so that threshold and filter conditions (size, diameter and perimeter of different spots) were established to train the application in a noise reduction method that determines the different positive activation signals: 0 for no signal, +1 for weak staining, +2 for moderate staining and +3 for strong staining. Immunohistochemistry for other proteins was conducted in a similar manner to the method described above using anti-p53 (Clone DO-7, Dako) and anti-WT1 (Clone WT1-562, Leica) antibodies, anti-TUBB4 (T7941, Sigma), anti-MYC (ab32072, Abcam), anti-EZH2 (Clone 6A10, Leica).

3. Results

3.1. Laparoscopy-guided Prospective Multi-region Sampling in an Ovarian Cancer Patient

We set out to discover early (root) driver mutations in a single high-grade serous ovarian tumor that may have contributed to its genesis.

We hypothesized that such early mutations would be present in all cancer cells irrespective of their location, time of presentation or biological behavior (e.g. chemotherapy response). To account for these three variables, we utilized intraoperative video recording to document the site of sampling before and after chemotherapy and at recurrence and to obtain a direct visual evaluation of chemotherapy response. LFR WGS technology enabled the accurate sequencing of very low (picogram) quantities of DNA (Peters et al., 2015) which is not currently possible by standard WGS or other methods of long fragment sequencing (Kuleshov et al., 2014). Since post-chemotherapy samples only had microscopic residual disease, using this technique on laser-capture microdissected cancer islets was essential.

We identified 750 unique mutations that were present in practically all tumor islet samples in our discovery set ($\geq 90\%$ of tumor islets) irrespective of space, tumor site or biological behavior. These mutations were extensively confirmed using standard WGS and targeted sequencing (Supplementary Fig. S2A). The variants included two synonymous mutations, eight non-synonymous mutations in protein-coding regions, and 740 mutations in non-coding regions (Supplementary Table 8). One mutation, detected in exon 8 of *TP53*, was shared between all tumor islets but was not present in normal stroma samples from patient 11152. This result was consistent with the previous finding by Ahmed et al. (2010) and Anon (2011) of ubiquitous *TP53* mutations in HGSOCS and has, therefore, validated our approach of discovering potential driver mutations.

3.2. Non-coding Mutations Cluster at Potential cis Regulatory Elements of Genetic Drivers of Stem Cell Differentiation

We mapped the 750 mutations to the nearest transcription start sites of genes and performed pathway-enrichment analyses of the identified genes as previously described (McLean et al., 2010). This analysis revealed five significantly enriched ontologies of biological processes (Supplementary Fig. S2B, Supplementary Tables 9 and 10) that were all related to embryonic and stem cell differentiation. Furthermore, the stem cell differentiation ontology was supported by 23 mutations that mapped to 15 genes (e.g., *SOX2*, *PAX7*, *WNT7A*). This finding was confirmed by several iterations of analysis (Supplementary Fig. S3) and multiple rounds of further sequencing (Supplementary Fig. S1A). The identification of *SOX2*, a key driver of stem cell differentiation (Takahashi and Yamanaka, 2006) that was recently implicated in skin cancer tumor initiation (Boumahdi et al., 2014), was particularly notable in our analysis because of the proximity of a total number of six mutations to the *SOX2* gene locus. We termed the nucleotides, in which the mutations mapping near *SOX2* occurred, BB1 to BB6.

3.3. The BB5 Mutation Occurred at the Pre-neoplastic Lesion of the Tumor and Marked a Region That Was Frequently Mutated in HGSOCS

We next performed deep-targeted sequencing of the 2 Mb region flanking the *SOX2* gene locus and BB1 to BB6 on tumor samples from 33 patients with HGSOCS, including case 11152 (Table 1, Supplementary Table 2), and identified 861 single nucleotide polymorphisms (SNPs) in this region (Supplementary Table 5) that were not previously reported in the 1000 Genomes Project (median = 21, range = 11 to 97 per cancer sample). Comparing the frequency of occurrence of these SNPs with those previously identified in the 1000 Genomes Project or with SNPs identified from whole genome sequencing analysis of a new set of 597 healthy elderly volunteers combined with deep whole genome re-sequencing of DNA from 436 individuals from the 1000 Genomes Project identified a 40 kb region flanking the BB5 nucleotide (Fig. 2A), that was significantly enriched in rare SNPs ($p < 0.01$, permutation analysis). These mutations significantly targeted biochemically active sites (implying that they were active regulatory elements) as determined by mining the ENCODE data for DNaseI hypersensitivity or transcription

Table 1

Patient baseline data for targeted sequencing analysis. The clinical details of patients who donated samples used for targeted sequencing are presented in this table.

Patients for targeted sequencing	49
Age at surgery	
Mean (median)	65.3 (66)
Min–max	40–82
<60 years	20.4%
>60 years	79.6%
Site of primary disease	
Ovary	29 (59%)
Tube	10 (20%)
Peritoneum	8 (16%)
Other ^a	2 (4%)
Tumor status	
Primary	19 (39%)
IDS	15 (31%)
Post-NACT	12 (24%)
Recurrence	3 (6%)
Biopsy site	
Ovary	17
Peritoneum ^b	15
Omentum	17
Other ^c	4
Normal tissue or blood ^d	24
Tumor stage ^e	
I–II	10 (20%)
III–IV	39 (80%)

Data are n (%). Disease stages I and II indicate early disease, stages III and IV indicate advanced disease. NACT = neoadjuvant chemotherapy. IDS = interval debulking surgery (e.g. after three or four cycles of NACT).

^a Ambivalent histology includes one case of tubal or ovarian origin and a case of ovarian or endometrial origin.

^b Peritoneum includes diaphragmatic peritoneum and small bowel serosa and mesentery.

^c Other biopsy sites include spleen, parasacral tumor and falciform ligament.

^d Normal tissue includes primary cell lines generated from normal fallopian tube epithelium.

^e For staging we used the FIGO classification valid at time of initial diagnosis.

factor binding activity (Supplementary Fig. S4). We therefore chose to characterize this 40 kb region, referred to as the BB5 region, further.

Sequencing results from tumor and normal tissue revealed that the BB5 region (chr3: 182,189,714–182,229,714) included 21 single nucleotide substitutions in HGSOCS from 16 patients (48.5% of all patients), of which nine variants from seven patients (21.2%) were somatic mutations (Supplementary Table 3). Independent validation by deep-targeted sequencing of a 1.6 Mb region flanking *SOX2* in tumor-normal tissue pairs of 16 additional patients with HGSOCS identified two further mutations in the BB5 region in two patients.

Examining the fallopian tube as a potential tissue of origin of HGSOCS (Crum, 2009; Lee et al., 2007; Kurman and Shih le, 2011; Karst et al., 2011) led to the identification of the p53 signature (Lee et al., 2007), indicated by p53 nuclear overexpression in fallopian tube epithelium (FTE), in patient 11152 (Fig. 2b–c). A subpopulation of the cells at the signature also expressed high levels of nuclear *SOX2* (Fig. 2d–e). Using a combination of deep-targeted sequencing, droplet-digital PCR and Sanger sequencing, we confirmed that both, the BB5 nucleotide and *TP53*, were already mutated in the p53 signature and the adjacent FTE (Fig. 2f–g). Importantly, in two additional patients for whom DNA extraction from the paraffin-embedded FTE was possible, we showed that the mutations occurring in the fully established HGSOCS were also present in the FTE (Supplementary Fig. S5).

3.4. Expansion of FTE Cells Strongly Expressing *SOX2* is a Feature of HGSOCS

The profound increase in the number of *SOX2*-expressing cells in the FTE of our index patient was in sharp contrast to the rare *SOX2* expression in the FTE of patients with benign conditions. Paradoxically, *SOX2*

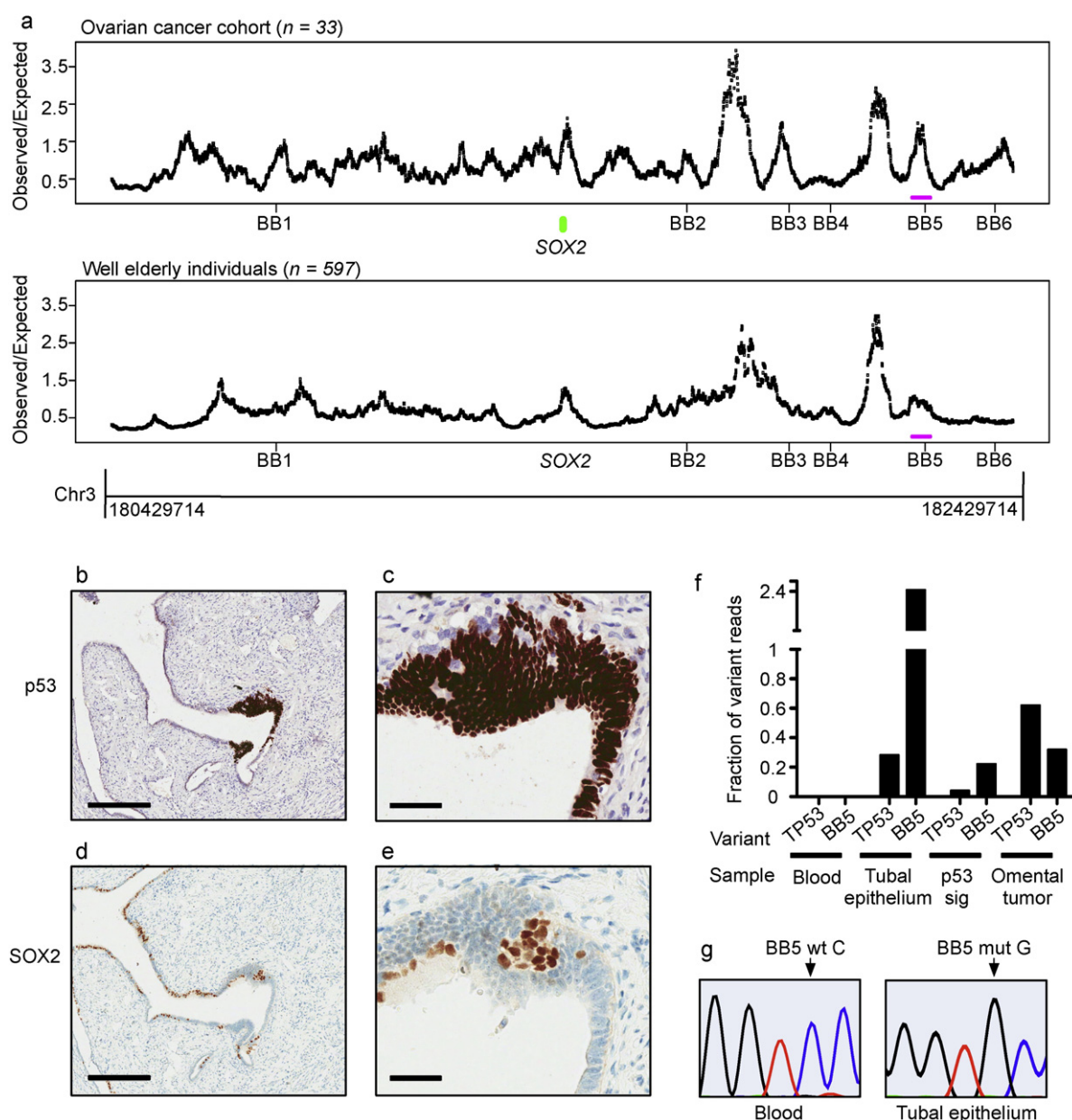


Fig. 2. A 40 kb region is frequently mutated in HGSOCs and consists of a distal SOX2 repressor. **a.** To determine whether any of the six mutations (BB1 to BB6) mapping near the SOX2 gene locus marked regions that were sites of frequent occurrence of private variants or mutations in HGSOCs, we performed deep-targeted sequencing of the 2 Mb region flanking SOX2 on 33 HGSOCs (Supplementary Table 2). A total of 861 single nucleotide substitutions (Supplementary Table 5) were identified that were not previously reported in the 1000 Genomes Project (median = 21, range = 11 to 97). Because functionally important genomic regions tend to be significantly less susceptible to genomic variation within a population, we determined whether the identified rare variants accumulated in specific areas within the 2 Mb region that were less susceptible to genomic alterations on a population scale. To test this hypothesis, we constructed overlapping moving windows of 40 kb size and compared the observed frequency of rare mutations (not previously described in the 1000 Genomes Project) in our group of patients. The expected frequency of SNPs in the same windows was based on 1000 sets of simulated cohorts of 33 individuals from the previously reported 1000 Genomes Project data. Upper panel: shown is the ratio of the observed number of variants in 40 kb “moving” windows in the cancer set to the expected number of variants in the equivalent windows based on 1000 permutations of simulated 1000 Genomes Project data. Based on this analysis, a peak observed/expected ratio (enrichment statistic) was identified in a 40 kb region flanking the BB5 nucleotide referred to as the BB5 region. Lower panel: To test whether this observation was higher than what would be expected by chance, we sequenced germline DNA from 597 healthy elderly volunteers and sequenced germline DNA from 436 individuals from the 1000 Genomes Project at higher depth. We then identified rare variants in the elderly set and repeated the above analysis. Comparing the enrichment statistic in the BB5 region in the cancer set to that obtained from 100 permutations of 33 individuals from the elderly set confirmed the significant enrichment of rare variants in the ovarian cancer set ($p < 0.01$). **b–e.** Immunohistochemical staining of p53 and SOX2 at low (**b** and **d**, scale bar = 300 μ m) and high (**c** and **e**, scale bar = 50 μ m) magnification in normal FTE and the p53 signature in patient 11152. Note the strong focal p53 staining at the multi-layered epithelium (p53 signature). **f.** The fraction of mutant reads relative to the total number of reads of the BB5 and TP53 mutations in germline DNA, the FTE, the p53 signature (p53 sig) and the tumor of the index case ($n = 4$ repeats). **g.** Sequencing trace indicating the BB5 mutation.

expression in tumor cells was almost absent (Supplementary Figs. S5 and S6). The increase in SOX2 expression in potential precursor cells is particularly intriguing since it is one of only four factors required for acquisition of stem cell properties by terminally differentiated cells (Takahashi and Yamanaka, 2006) with important implication for acquisition of pluripotency, the ability of a cell to differentiate into different cell types, and susceptibility to tumor formation.

We next compared SOX2 expression in the nuclei of normal FTE of patients with benign conditions and patients with endometrial cancer (age-matched cases) and the normal FTE and corresponding ovarian tumors from patients diagnosed with HGSOC (Table 2, Supplementary Table 7). Automated image analysis of these samples demonstrated that the median SOX2 expression (intensity score of 3+) in the normal appearing FTE from patients with HGSOCs was 17-fold higher than that

Table 2

Patient characteristics for immunohistochemistry studies. The clinical details of patients who donated the samples used for SOX2 immunohistochemistry are presented.

Patients for IHC staining <i>n</i> = 279	Benign	HGSOC	BRCA mutation carrier
Patient cohort	108	100	71
Discovery cohort	16	22 ^a	
Validation cohort	92	78	71
Age at surgery			
Mean (median)	48.0 (48)	62.7 (63)	50.7 (48)
Min–max	28–81	32–94	29–81
Site of primary disease			
Ovary		79 (79%)	
Tube		10 (10%)	
Peritoneum		5 (5%)	
Other ^b		6 (6%)	
Tumor status			
Primary		60 (60%)	
IDS		15 (15%)	
Post-NACT		19 (19%)	
Tumor stage ^c			
I–II		26 (26%)	
III–IV		69 (69%)	

Data are *n* (%) or mean (median), unless otherwise indicated. Disease stages I and II indicate early disease, stages III and IV indicate advanced disease. IDS = Interval debulking surgery (e.g. after three or four cycles of NACT). NACT = neoadjuvant chemotherapy.

^a These patients were also included in targeted sequencing study.

^b Ambivalent histology cohort includes four cases of ovarian or tubal origin and two cases of ovarian or peritoneal origin.

^c For staging we used the FIGO classification valid at time of initial diagnosis.

in the normal FTE from patients with benign conditions ($p < 2^{-16}$, one-way analysis of variance [ANOVA] followed by Tukey's test). The median SOX2 expression significantly decreased in the corresponding, fully established ovarian tumors ($p < 2^{-16}$, one-way ANOVA followed by Tukey's test, Fig. 3a–b). We validated this observation on an independent set of fallopian tubes from 88 women with either HGSOCs ($n = 42$) or benign gynecological conditions ($n = 46$) and confirmed that a significant expansion of cells strongly expressing SOX2 was present in the benign FTE of women with HGSOCs ($p < 2^{-16}$, one-way ANOVA followed by Tukey's test, Fig. 3c). The magnitude of SOX2 overexpression was not significantly influenced by tumor stage (early versus late) or by whether or not the patient had chemotherapy prior to surgical excision (data not shown). In addition, fitting a binomial general linear model on the entire dataset revealed that strong SOX2 expression accurately predicted whether the fallopian tube was from an HGSOC patient or from a patient with benign pathology ($p = 9.95^{-8}$, logistic regression model). Fitting a receiver operating characteristic (ROC) curve to the data from the logistic regression analysis confirmed the high predictive power of strong SOX2 expression (Fig. 3d, area under the curve [AUC] = 0.87). This ubiquitous feature of the FTE of patients with HGSOCs has not been previously described. Furthermore, only eight out of 28 cases stained for p53 also showed a p53 signature indicating that SOX2 overexpression may occur prior to p53 dysregulation.

To test whether the expansion of SOX2-expressing cells occurred prior to HGSOC development, we analyzed the fallopian tubes of 48 women at high risk of developing ovarian cancer because they were BRCA1 or BRCA2 gene mutation carriers and therefore underwent prophylactic salpingo-oophorectomy (Fig. 3c, Table 2). This analysis showed that a significant expansion of cells strongly expressing SOX2 was present in the benign FTE of women who were BRCA1 or BRCA2 mutation carriers ($p < 0.001$, one-way ANOVA followed by Tukey's test).

Finally, to confirm the above findings, we repeated the analysis on an additional cohort of 98 patients from three independent institutes (Fig. 3e, Table 2 and Supplementary Table 7). This analysis further confirmed that a significant expansion of cells strongly expressing SOX2 occurred in the benign FTE of women who are BRCA1 or BRCA2 mutation carriers and in women who had HGSOCs ($p = 0.018$ and $p < 0.001$, respectively,

one-way ANOVA followed by Tukey's test). Expression of SOX2 in the normal FTE of BRCA gene mutation carriers was independent of age and whether the mutation was present in the BRCA1 or BRCA2 (data not shown). History of prior breast cancer was available for 43 cases who were either BRCA1 or BRCA2 mutation carriers. Of those only 15 patients had history of prior breast cancer. There was no significant difference in SOX2 expression between cases that either had or did not have breast cancer.

3.5. The BB5 Region Is a Repressor of SOX2 Expression

Specific sequences of non-coding regions in the genome are capable of regulating the expression of neighboring genes. Several lines of evidence strongly suggested that the BB5 region acted as a repressor of SOX2 expression. First, the BB5 mutation occurred in a 12-nucleotide DNA sequence (Supplementary Fig. S7A) that matched the previously reported regulatory motif model UW.0169 that was discovered by genomic DNaseI footprinting (Neph et al., 2012). We used an in-vivo reporter assay to test whether this motif changed the expression of a fluorescent protein and, therefore, was potentially functional (Supplementary Fig. S7B–E). Such standard reporter assays are traditionally performed in cultured cells. In this study we conducted the assay in live chicken embryos to gain additional information about the time and location of the activity of the BB5 region in regulating fluorescent protein (the reporter) expression. This analysis confirmed that the motif consisting of the BB5 nucleotide acted as a repressor and that mutation of the BB5 nucleotide significantly relieved this repressor activity (Supplementary Fig. S7B–E). Importantly, the protein fluorescence partially co-localized with cells that strongly expressed endogenous Sox2 in the chicken embryo and, therefore, implied that the element was plausible for regulating Sox2 expression. Second, we performed CRISPR-Cas9-mediated genome deletion, a method to remove a number of nucleotides from a target sequence in the genome, and deleted multiple essential nucleotides included in the motif model UW.0169 (Supplementary Fig. S7F) (Ran et al., 2013) in a human cell model. This induced a significant reduction in H3K27 acetylation, an indicator of open chromatin, and a significant and sustained increase in SOX2 expression in human cells (Supplementary Fig. S7G, Supplementary Table 6). Thus, this unequivocally confirmed that the element acted as a repressor of SOX2 expression in a human cell model. Third, mining the ENCODE data for potential transcription factor binding at this element revealed several factors that are known to be involved in regulating the stem cell state, such as YY1, NANOG, OCT4 and EZH2 (Supplementary Fig. S7F). Since EZH2, the catalytic subunit of the Polycomb repressor complex 2, is known to be involved in repressing gene expression by inducing H3K27 trimethylation at regulatory elements, we performed ChIP-sequencing on primary cultured fallopian tube epithelial cells to test whether EZH2 occupied the BB5 element or other elements that were sites of somatic mutations in other HGSOCs. This analysis confirmed that EZH2 occupied the BB5 element and several others (Supplementary Fig. S8A, Supplementary Table 11). In addition, the enzymatic inhibition of EZH2 resulted in a significant reduction in H3K27 tri-methylation (Supplementary Fig. S8B) indicating that the localization of the EZH2 at the BB5 element was indeed functional. In summary, the above data confirmed that the BB5 region is a repressor of SOX2 expression and, when mutated, releases SOX2 repression.

As previously mentioned, SOX2 is part of the four essential stem cells transcription factors (SOX2, OCT4, c-Myc, KLF4), also known as Yamanaka factors (Takahashi et al., 2007), which are highly expressed in embryonic stem cells and that are sufficient to induce pluripotency in human somatic cells. We, therefore, investigated the expression of the Yamanaka factors in the FTE and observed a mutually exclusive expression pattern of SOX2 and MYC. In the FTE of benign cases, MYC appeared to be exclusively and ubiquitously expressed in non-ciliated cells whilst SOX2 was only expressed at low levels in a minor population of

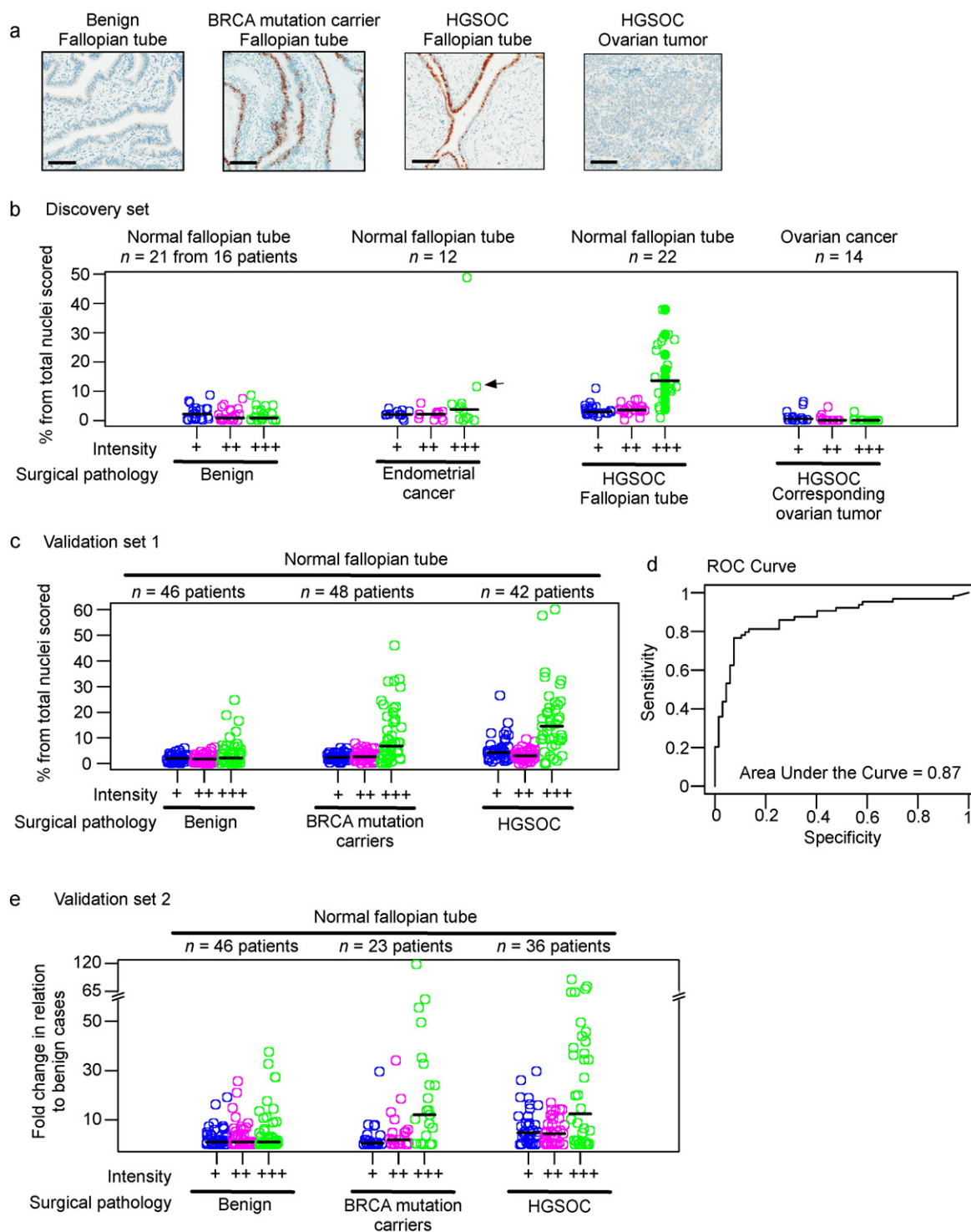


Fig. 3. Overexpression of SOX2 is a biomarker for pre-neoplastic detection of HGSOC. **a.** Representative SOX2 immunohistochemistry images for the FTE of the indicated samples are presented. Scale bars = 100 μ m. **b.** Percentages (y-axis), median percentage (horizontal black bars) and intensity (x-axis) of SOX2 staining in the normal FTE of women with benign conditions, endometrial cancer or HGSOCs and in the paired HGSOC tumors. Solid circles represent the FTE from cancers harboring rare variants and mutations in the BB5 region. The black arrow indicates a case of high-grade serous endometrial cancer. **c.** Power calculations were used to determine the required case number for the validation set based on the data from the discovery set (see Supplementary methods). Data from an independent set of 88 cases as well as an additional 48 BRCA1 or BRCA2 mutation carriers who underwent prophylactic excision of the fallopian tubes are presented. **d.** The receiver operating characteristics (ROC) curve is presented for the combined data presented in panels b and c. **e.** Data from a further independent set of cases from a second institute (CNIO) are presented. Data are presented as fold change relative to benign cases. A y-axis break was added to facilitate comparison with the data presented in b and c.

ciliated cells (Fig. 4). In contrast, in the FTE from HGSOC cases MYC maintained a similar expression pattern in non-ciliated cells, whereas SOX2 was strongly expressed in the majority of ciliated cells (Fig. 4, Supplementary Fig. S9A–B). Furthermore, the analysis of 209 HGSOCs revealed that MYC over-expression was accompanied by the loss of

SOX2 expression and vice versa ($p < 0.001$, Fisher Exact test, Supplementary Fig. S9C). Thus, SOX2 may repress MYC in differentiating FTE cells and in a fraction of SOX2-expressing tumors. Consistent with our proposed role of SOX2 in FTE cells, its ectopic expression significantly reduced expression of MYC and PAX8, an important transcription factor

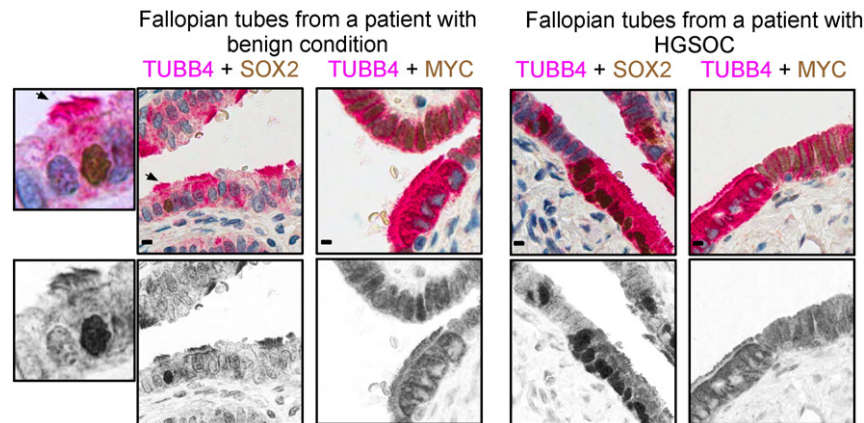


Fig. 4. SOX2 represses MYC and PAX8 expression at the secretory-ciliated FTE cell interface. Representative immunohistochemistry images of FTE double stained with the indicated antibodies ($n = 10$). TUBB4 was used as a marker for ciliated cells. Scale bars = 10 μ m. Arrows indicate cells that are presented in higher magnifications.

during embryogenesis of Müllerian tissues, in primary cultured FTE cells and SKOV3 ovarian cancer cells to undetectable levels by immunostaining (Supplementary Fig. S10). These findings are consistent with the known role of SOX2 in inducing pluripotency and may have important implications in understanding the genesis of epithelial ovarian cancers.

4. Discussion

Less than a third of ovarian cancer patients are diagnosed at an early, hence potentially curable, stage. The currently available detection options are patient driven symptom awareness and screening using combination of a serum marker (CA125) and imaging by ultrasound methods. Advances have been made to divert from the single threshold rule for CA125 to more complex algorithms in order to improve performance and facilitate early treatment. Jacobs et al. (2016) recently reported results of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), a study in which over 200,000 postmenopausal women were enrolled to assess whether screening is effective for improving early ovarian cancer diagnosis and survival. Results indicated a limited mortality reduction in the primary analysis (Jacobs et al., 2016). This, again, underscores the need for effective biomarkers for early detection.

Recent work strongly suggests that the FTE is a common site of a large fraction of high-grade serous pelvic cancers (Perets et al., 2013). Inducing mutations in *TP53*, *BRCA1-BRCA2*, and *PTEN* in secretory cells in the FTE of mice resulted in the development of HGSOcs (Perets et al., 2013). However, there continues to be a need for novel molecular markers of premalignant transformation that could be used for early detection. Current screening methods only enable the detection of established invasive ovarian cancers. Mathematical modeling of tumor occurrence in *BRCA1-BRCA2* mutations carriers suggests that tumors remain occult for up to four years (Brown and Palmer, 2009). This duration represents an unutilized opportunity for premalignant or early stage disease detection because of the limitations of existing markers. In addition, patients with a family history of ovarian cancer who don't have affected living relatives are not eligible for *BRCA1-BRCA2* mutation testing. Furthermore, in many cases of familial ovarian malignancies *BRCA* testing is negative. Subsequently, these patients are ineligible for risk-reducing salpingo-oophorectomy.

Lack of specificity of *TP53* overexpression (the so called p53 signature) in the fallopian tube (Lee et al., 2007) undermines its potential use in early detection. In contrast, we have shown that SOX2 overexpression has high specificity and sensitivity as evidenced by the high AUC in our ROC analysis. In addition, unlike the p53 signature which only involves a limited number of cells, SOX2 overexpression occupies a large surface area of the fallopian tube. This makes SOX2 overexpression amenable for detection. Safe screening methods that exploit such

feature would need to be developed in both *BRCA* mutation carriers and patients with HGSOc. The use of cell-penetrating peptides or antibodies that specifically recognize SOX2 could be labeled with PET imaging tracers to enable early, non-invasive detection in patients. In addition, fallopian tube catheterization for obtaining cytobrush specimens for quantifying SOX2 expression is also possible. Such methods could be tested in clinical trials for early detection in young women with *BRCA1-BRCA2* mutations who do not wish to undergo prophylactic excision of both tubes and ovaries. Finally, as SOX2 is a potent transcription factor, it is highly likely that its aberrant expression is associated with co-expression of secreted proteins. The future discovery of such proteins could enable biomarker screening using blood sampling.

In summary, in this study we demonstrated that SOX2 overexpression occurs in a fraction of women with *BRCA1* and *BRCA2* mutations prior to ovarian cancer initiation and in the majority of patients with HGSOcs irrespective of tumor stage. These findings could be exploited for filling the current gap in early detection strategies for ovarian cancer. We believe that this is the first report of the expansion of SOX2-expressing cells in the FTE of HGSOcs. This finding has important implications, as it provides a potentially powerful tool for screening for HGSOcs. Utilizing our findings as potential biomarker should take high priority.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ebiom.2016.06.048>.

Funding Sources

This work is funded by the Medical Research Council (H8RSRS00), Ovarian Cancer Action (HER00070), the Oxford Biomedical Research Centre, the National Institute for Health Research (HJRWAC05) and the Experimental Cancer Medicine Centre. MJG is recipient of a research contract from the Instituto de Salud Carlos III of the Ministerio Español de Sanidad y Consumo (Miguel Servet tipo II Program, CP11 13-00047). C.Y. acknowledges the support of an MRC New Investigator Research Grant (Ref No. MR-L001411-1) and the Wellcome Trust Core Award Grant Number 090532-Z-09-Z.

Conflicts of Interests

R.T., B.A.P., R.D., D.M.H., R.Y.Z., R.K., and K.D. own stock in Complete Genomics, Inc. All other authors have no conflict of interest to declare.

Contributions

A.A.A. conceived the project. All other authors designed and conducted experiments. A.A.A., K.H., F.M., S.H.G., and C.Y. wrote the manuscript. All authors contributed to the final manuscript.

Acknowledgements

We thank the Gynecological Oncology Multidisciplinary Team at Oxford, the Early Phase Clinical Trials Unit and the Oncology Clinical Trials Office at Oxford for help in recruiting patients to clinical trials. We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust grant reference 090532-Z-09-Z and MRC Hub grant G0900747 91070) for their help with targeted sequencing and Stephen Hague for assistance with droplet digital PCR. We thank the Oxford Centre for Histopathology Research for their technical work and Hayley Davis for assisting with laser capture microdissection. We would like to acknowledge the support of all Complete Genomics employees who worked on the generation of libraries, reagents, and whole genome sequencing data for this project. We thank the CNIO Histopathology Core Unit for technical work. We thank the biobanks in CNIO (Ref. B.000848), MD Anderson (Ref. B.0000745) and IRBLleida (Ref. B.0000682) and the Spanish National Biobank Network Platform (Instituto de Salud Carlos III) for coordinating sample and clinical data retrieval.

References

- Abecasis, G.R., Auton, A., Brooks, L.D., Depristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., McVean, G.A., 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature* 491, 56–65.
- Ahmed, A.A., Etemadmoghadam, D., Temple, J., Lynch, A.G., Riad, M., Sharma, R., Stewart, C., Feraday, S., CALDAS, C., Defazio, A., Bowtell, D., Brenton, J.D., 2010. Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J. Pathol.* 221, 49–56.
- Anon, 2011. Integrated genomic analyses of ovarian carcinoma. *Nature* 474, 609–615.
- Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., Snyder, M., 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
- Besnier, C., Takeuchi, Y., Towers, G., 2002. Restriction of lentivirus in monkeys. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11920–11925.
- Betancur, P., Bronner-Fraser, M., Sauka-Spengler, T., 2010. Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. *Proc. Natl. Acad. Sci. U. S. A.* 107, 3570–3575.
- Boumahdi, S., Driessens, G., Lapouge, G., Rorive, S., Nassar, D., Le Mercier, M., Delatte, B., Caauwe, A., Lenglez, S., Nkusi, E., Brohee, S., Salmon, I., Dubois, C., Del Marmol, V., Fuks, F., Beck, B., Blanpain, C., 2014. SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature* 511, 246–250.
- Brown, P.O., Palmer, C., 2009. The preclinical natural history of serous ovarian cancer: defining the target for early detection. *PLoS Med.* 6, e1000114.
- Carnevali, P., Baccash, J., Halpern, A.L., Nazarenko, I., Nilsen, G.B., Pant, K.P., Ebert, J.C., Brownley, A., Morenzoni, M., Karpinchyk, V., Martin, B., Ballinger, D.G., Drmanac, R., 2012. Computational techniques for human genome resequencing using mated gapped reads. *J. Comput. Biol.* 19, 279–292.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823.
- Crum, C.P., 2009. Intercepting pelvic cancer in the distal fallopian tube: theories and realities. *Mol. Oncol.* 3, 165–170.
- Dean, F.B., Hosono, S., Fang, L., Wu, X., Faruqi, A.F., Bray-Ward, P., Sun, Z., Zong, Q., Du, Y., Du, J., Driscoll, M., Song, W., Kingsmore, S.F., Egholm, M., Lasken, R.S., 2002. Comprehensive human genome amplification using multiple displacement amplification. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5261–5266.
- Drescher, C.W., Shah, C., Thorpe, J., O'Brian, K., Anderson, G.L., Berg, C.D., Urban, N., McIntosh, M.W., 2013. Longitudinal screening algorithm that incorporates change over time in CA125 levels identifies ovarian cancer earlier than a single-threshold rule. *J. Clin. Oncol.* 31, 387–392.
- Drmanac, R., Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kermani, B.G., Carnevali, P., Nazarenko, I., Nilsen, G.B., Yeung, G., Dahl, F., Fernandez, A., Staker, B., Pant, K.P., Borchert, A.P., Brownley, A., Cedenio, R., Chen, L., Chernikoff, D., Cheung, A., Chirita, R., Curson, B., Ebert, J.C., Hacker, C.R., Hartlage, R., Hauser, B., Huang, S., Jiang, Y., Karpinchyk, V., Koenig, M., KONG, C., Landers, T., Le, C., Liu, J., McBride, C.E., Morenzoni, M., Morey, R.E., Mutch, K., Perazich, H., Perry, K., Peters, B.A., Peterson, J., Pethiyagoda, C.L., Pothuraju, K., Richter, C., Rosenbaum, A.M., Roy, S., Shafit, J., Sharanovich, U., Shannon, K.W., Sheppy, C.G., Sun, M., Thakuria, J.V., Tran, A., Vu, D., Zaranek, A.W., Wu, X., Drmanac, S., Oliphant, A.R., Banyai, W.C., Martin, B., Ballinger, D.G., Church, G.M., Reid, C.A., 2010. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 327, 78–81.
- Feng, J., Liu, T., Qin, B., Zhang, Y., Liu, X.S., 2012. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* 7, 1728–1740.
- Hamburger, V., Hamilton, H.L., 1992. A series of normal stages in the development of the chick embryo. 1951. *Dev. Dyn.* 195, 231–272.
- Jacobs, I.J., Menon, U., Ryan, A., Gentry-Maharaj, A., Burnell, M., Kalsi, J.K., Amso, N.N., Apostolidou, S., Benjamin, E., Cruickshank, D., Crump, D.N., Davies, S.K., Dawnay, A., Dobbs, S., Fletcher, G., Ford, J., Godfrey, K., Gunu, R., Habib, M., Hallett, R., Herod, J., Jenkins, H., Karpinskyj, C., Leeson, S., Lewis, S.J., Liston, W.R., Lopes, A., Mould, T., Murdoch, J., Oram, D., Rabideau, D.J., Reynolds, K., Scott, I., Seif, M.W., Sharma, A., Singh, N., Taylor, J., Warburton, F., Widschwendter, M., Williamson, K., Woolas, R., Fallowfield, L., McGuire, A.J., Campbell, S., Parmar, M., Skates, S.J., 2016. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKTOCS): a randomised controlled trial. *Lancet* 387, 945–956.
- Karst, A.M., Levanon, K., Drapkin, R., 2011. Modeling high-grade serous ovarian carcinogenesis from the fallopian tube. *Proc. Natl. Acad. Sci. U. S. A.* 108, 7547–7552.
- Kuleshov, V., Xie, D., Chen, R., Pushkarev, D., Ma, Z., Blauwkamp, T., Kertesz, M., Snyder, M., 2014. Whole-genome haplotyping using long reads and statistical methods. *Nat. Biotechnol.* 32, 261–266.
- Kurman, R.J., Shih, Ie, M., 2011. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer—shifting the paradigm. *Hum. Pathol.* 42, 918–931.
- Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., Chen, Y., Desalvo, G., Epstein, C., Fisher-Aylor, K.I., Euskirchen, G., Gerstein, M., Gertz, J., Hartemink, A.J., Hoffman, M.M., Iyer, V.R., Jung, Y.L., Karmakar, S., Kellis, M., Kharchenko, P.V., Li, Q., Liu, T., Liu, X.S., Ma, L., Milosavljevic, A., Myers, R.M., Park, P.J., Pazin, M.J., Perry, M.D., Raha, D., Reddy, T.E., Rozowsky, J., Shores, N., Sidow, A., Slattery, M., Stamatoyannopoulos, J.A., Tolstorukov, M.Y., White, K.P., Xi, S., Farnham, P.J., Lieb, J.D., Wold, B.J., Snyder, M., 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* 22, 1813–1831.
- Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Lee, Y., Miron, A., Drapkin, R., Nucci, M.R., Medeiros, F., Saleemuddin, A., Garber, J., Birch, C., Mou, H., Gordon, R.W., Cramer, D.W., McKeon, F.D., Crum, C.P., 2007. A candidate precursor to serous carcinoma that originates in the distal fallopian tube. *J. Pathol.* 211, 26–35.
- Lunter, G., Goodson, M., 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* 21, 936–939.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaer, B.T., Lowe, C.B., Wenger, A.M., Bejerano, G., 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* 28, 495–501.
- Menon, U., Ryan, A., Kalsi, J., Gentry-Maharaj, A., Dawnay, A., Habib, M., Apostolidou, S., Singh, N., Benjamin, E., Burnell, M., Davies, S., Sharma, A., Gunu, R., Godfrey, K., Lopes, A., Oram, D., Herod, J., Williamson, K., Seif, M.W., Jenkins, H., Mould, T., Woolas, R., Murdoch, J.B., Dobbs, S., Amso, N.N., Leeson, S., Cruickshank, D., Scott, I., Fallowfield, L., Widschwendter, M., Reynolds, K., McGuire, A., Campbell, S., Parmar, M., Skates, S.J., Jacobs, I., 2015. Risk algorithm using serial biomarker measurements doubles the number of screen-detected cancers compared with a single-threshold rule in the united kingdom collaborative trial of ovarian cancer screening. *J. Clin. Oncol.* 33, 2062–2071.
- Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., Maurano, M.T., Humbert, R., Rynes, E., Wang, H., Vong, S., Lee, K., Bates, D., Diegel, M., Roach, V., Dunn, D., Neri, J., Schafer, A., Hansen, R.S., Kutayavin, T., Giste, E., Weaver, M., Canfield, T., Sabo, P., Zhang, M., Balasundaram, G., Byron, R., Macoss, M.J., Akey, J.M., Bender, M.A., Groudine, M., Kaul, R., Stamatoyannopoulos, J.A., 2012. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489, 83–90.
- Perets, R., Wyant, G.A., Muto, K.W., Bijron, J.G., Poole, B.B., Chin, K.T., Chen, J.Y., Ohman, A.W., Stepule, C.D., Kwak, S., Karst, A.M., Hirsch, M.S., Setlur, S.R., Crum, C.P., Dinulescu, D.M., DRAPKIN, R., 2013. Transformation of the fallopian tube secretory epithelium leads to high-grade serous ovarian cancer in Brca1;Pten models. *Cancer Cell* 24, 751–765.
- Peters, B.A., Kermani, B.G., Sparks, A.B., Alferov, O., Hong, P., Alexeev, A., Jiang, Y., Dahl, F., Tang, Y.T., Haas, J., Robasky, K., Zaranek, A.W., Lee, J.H., Ball, M.P., Peterson, J.E., Perazich, H., Yeung, G., Liu, J., Chen, L., Kennemer, M.L., Pothuraju, K., Konvicka, K., Tsoupo-Gitnikov, M., Pant, K.P., Ebert, J.C., Nilsen, G.B., Baccash, J., Halpern, A.L., Church, G.M., Drmanac, R., 2012. Accurate whole-genome sequencing and haplotyping from 10 to 20 human cells. *Nature* 487, 190–195.
- Peters, B.A., Kermani, B.G., Alferov, O., Agarwal, M.R., McElwain, M.A., Gulbahce, N., Hayden, D.M., Tang, Y.T., Zhang, R.Y., Tearle, R., Crain, B., Prates, R., Berkeley, A., Munne, S., Drmanac, R., 2015. Detection and phasing of single base de novo mutations in biopsies from human in vitro fertilized embryos by advanced whole-genome sequencing. *Genome Res.* 25, 426–434.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308.
- Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S.R., Consortium, W. G. S., Wilkie, A.O., McVean, G., Lunter, G., 2014. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat. Genet.* 46, 912–918.
- Sauka-Spengler, T., Barembaum, M., 2008. Gain- and loss-of-function approaches in the chick embryo. *Methods Cell Biol.* 87, 237–256.
- Simoes-Costa, M.S., McKeown, S.J., Tan-Cabugao, J., Sauka-Spengler, T., Bronner, M.E., 2012. Dynamic and differential regulation of stem cell factor FoxD3 in the neural crest is encrypted in the genome. *PLoS Genet.* 8, e1003142.
- Takahashi, K., Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872.
- Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., Garg, K., John, S., Sandstrom, R., Bates, D., Boatman, L., Canfield, T.K., Diegel, M., Dunn, D., Ebersol, A.K., Frum, T., Giste, E., Johnson, A.K., Johnson, E.M., Kutayavin, T., Lajoie, B., Lee, B.K., Lee, K., London, D., Lotakis, D., Neph, S., Neri, F., Nguyen, E.D., Qu, H., Reynolds, A.P., Roach, V., Safi, A., Sanchez, M.E., Sanyal, A., Shafer, A., Simon, J.M., Song, L., Vong, S., Weaver, M., Yan,

- Y., Zhang, Z., Lenhard, B., Tewari, M., Dorschner, M.O., Hansen, R.S., Navas, P.A., Stamatoyannopoulos, G., Iyer, V.R., Lieb, J.D., Sunyaev, S.R., Akey, J.M., Sabo, P.J., Kaul, R., Furey, T.S., Dekker, J., Crawford, G.E., Stamatoyannopoulos, J.A., 2012. The accessible chromatin landscape of the human genome. *Nature* 489, 75–82.
- Yang, X., Boehm, J.S., Salehi-Ashtiani, K., Hao, T., Shen, Y., Lubonja, R., Thomas, S.R., Alkan, O., Bhimdi, T., Green, T.M., Johannessen, C.M., Silver, S.J., Nguyen, C., Murray, R.R., Hieronymus, H., Balcha, D., Fan, C., Lin, C., Ghamsari, L., Vidal, M., Hahn, W.C., Hill, D.E., Root, D.E., 2011. A public genome-scale lentiviral expression library of human ORFs. *Nat. Methods* 8, 659–661.
- Yip, K.Y., Cheng, C., Bhardwaj, N., Brown, J.B., Leng, J., Kundaje, A., Rozowsky, J., Birney, E., Bickel, P., Snyder, M., Gerstein, M., 2012. Classification of human genomic regions based on experimentally determined binding sites of more than 100 transcription-related factors. *Genome Biol.* 13, R48.